CHARACTER AND PERSONALITY

An International Psychological Quarterly

Vol. VI DECEMBER, 1937 No. 2

CONTENTS

PAGE

WINIFRED BENT JOHNSON (Stanford University):
"Euphoric and Depressed Moods in Normal Subjects, Part I" 79

P. E. VERNON (Jordanhill Training Center, Glasgow):
"The Stanford-Binet Test as a Psychometric Method"... 99

RAYMOND B. CATTELL (Psychological Clinic, Leicester, England):
"Measurement Versus Intuition in Applied Psychology".. 114

WILLIAM STERN (Duke University):
"Cloud Pictures: A New Method for Testing Imagination" 132

FRANZISKA BAUMGARTEN (University of Bern):
"Character Traits Derived from Biographies"..... 147

DONALD W. MACKINNON (Bryn Mawr College):
"Trends in the Study of Personality (As Revealed in the 1937 Meeting of the American Psychological Association)" 150

Book Reviews 156

Books Received 161

DUKE UNIVERSITY PRESS CENTENNIAL PRIZE

In connection with the Centennial of the Origins of Trinity College, now a part of Duke University, which is to be celebrated during the academic year 1938-1939, the Duke University Press announces a prize of $1,500 for a scholarly manuscript in the fields of the social, literary, or artistic history of the United States, to be awarded on March 1, 1939.

Conditions

1. The manuscript must be in the fields mentioned above, must be scholarly in nature, and must be of a length no less than fifty thousand (50,000) words.
2. All manuscripts must be typed, signed by a nom-de-plume, and must be submitted no later than October 1, 1938.
3. A sealed envelope containing the author's name and address within and bearing his nom-de-plume on the outside must accompany each manuscript.
4. Any scholar in the United States excepting members of the faculty of Duke University is eligible to compete.
5. The winning manuscript will be published by the Duke University Press, and its author will receive the regular royalty of ten per cent in addition to the prize.
6. A committee of the Duke University faculty will select from all manuscripts entered the three which in their opinion seem most worthy of the prize.
7. A committee of three distinguished scholars not connected with Duke University will make the final selection of the prize-winning manuscript.
8. All manuscripts should be addressed to the Duke University Press, and authors are requested to keep copies of the manuscripts submitted since there can be no positive guarantee against loss.

* * * * * *

Duke Faculty Committee

William T. Laprade
(Chairman)

Charles Roberts Anderson
Clarence Gohdes
Calvin Bryce Hoover
Jay Broadus Hubbell
Richard Harrison Shryock
Charles Sackett Sydnor

Final Judges

Merle Eugene Curti
(Smith College)

Ralph Leslie Rusk
(Columbia University)

Arthur Meier Schlesinger
(Harvard University)

EUPHORIC AND DEPRESSED MOODS IN NORMAL SUBJECTS*

PART I

WINIFRED BENT JOHNSON
Stanford University

This study is an attempt to discover the behavioral correlates and the temporal patterns of normal shifts in elation-depression. In particular the author wished to test the hypothesis that the abnormal affective states in psychotics are only exaggerations of normal mood swings and that a continuum exists from normal to abnormal.

The subjects employed were thirty university women, all resident members of a sorority of the University of California. Each subject agreed to furnish daily self-rating on mood status and to serve as subject in two experimental sessions of one and one-half hours each.

The data to be presented fall in three groups:

1. Temporal patterns of euphoria and depression. The data for this section are estimates of the current day's feeling tone on an eleven-point scale over periods of sixty-five to ninety consecutive days. Samples of mood patterns are presented graphically. A composite graph is also discussed.

2. Behavior characteristic of euphoric moods contrasted with that of depressed moods. These data consist of quantitative measurements of objective behavior in two opposed moods. The aspects of behavior measured were chosen from those commonly found to differentiate elated from depressed psychopaths.

3. Subjective accompaniments of euphoric moods contrasted with those of depressed moods. These data consist of self-ratings on emotional variables and were taken at two sessions, one in each extreme mood. The emotional variables rated were chosen from those appearing to differentiate elated from depressed psychopaths.

* The author wishes to acknowledge the assistance of Dr. Robert Choate Tryon and other psychologists of the staff of the University of California, and of Dr. Lewis Madison Terman of Stanford University.

The theoretical bearing of these data will be discussed in a final section.

A. Temporal Patterns of Mood

Every evening during the course of the experiment, the experimenter had each subject rate herself for euphoria-depression by checking an eleven-point scale. Ratings were taken immediately before dinner. In case of absence from dinner no rating was accepted unless made during the evening of the current day.

Definition of the scale points between the extremes was not made; and extremes were defined only in relation to the self, as "I almost never feel more elated," or "as depressed as I practically ever feel." The logic of this was as follows: every intelligent individual is able to differentiate one of his own definitely "low" moods from one of his own definitely "high" moods. Detailed definition of points is, therefore, unnecessary. More than this, an attempt to define scale points in absolute terms would be unjustified, because the specific emotional components of varying degrees of elation-depression are unknown.

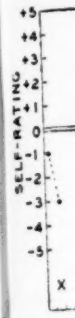
In this study it has not been assumed that steps on the scale, as conceived by one subject, are equal or similar to those steps as conceived by another subject. This statement is of the utmost importance for an understanding of the aims of the study. The writer attempts to discover some of the changes which occur within the individual as mood changes. Inter-personality comparisons are valid for only a limited portion of the data and are largely irrelevant to the goals of the study. These goals are to discover the nature and temporal pattern of mood changes by comparing the individual with himself.

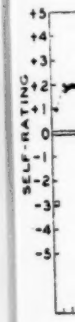
Examples of Mood Patterns

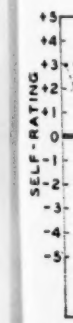
Figures I-VI show mood changes reported by six unselected subjects during the entire period of the study. Arrows indicate days when experimental data were gathered. Bars indicate days on which menstruation was reported, and crosses indicate other types of illness.

Features of the Temporal Patterns of Mood

The tendency toward rhythm in our data is not as marked as Hersey's (8) statements regarding his longer observations might lead one to expect, but it appears greater than reaction to chance variations in the environment would give. Rhythm is certainly







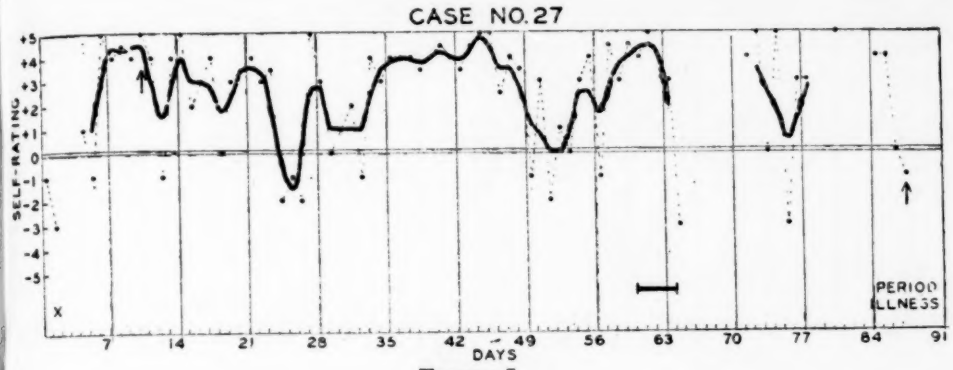


FIGURE I

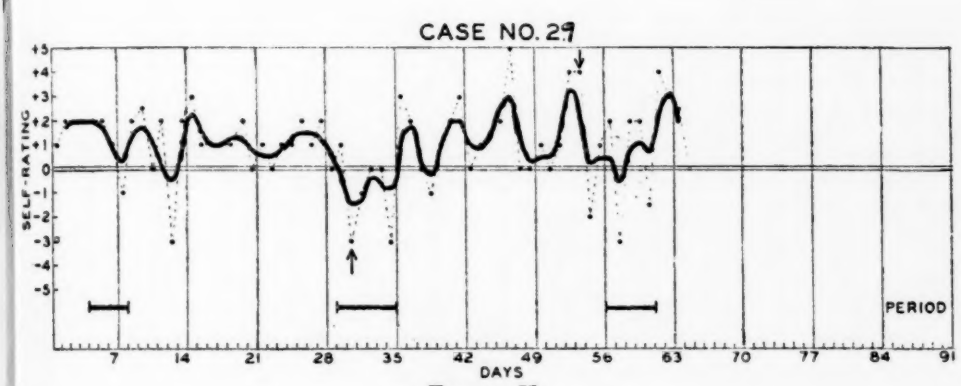


FIGURE II

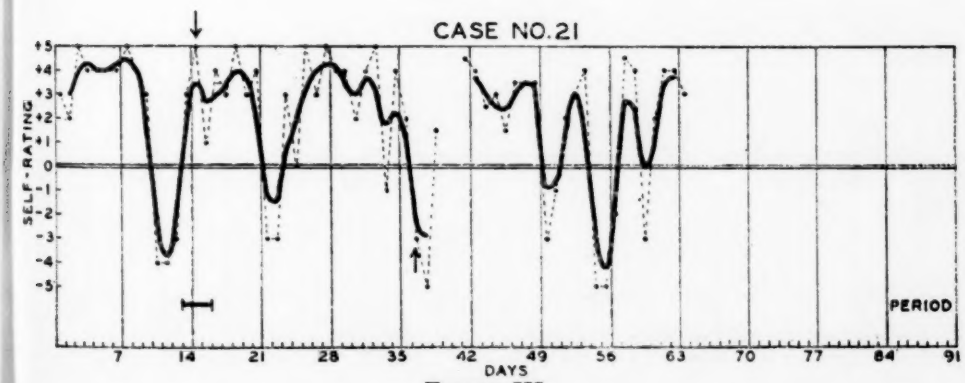


FIGURE III

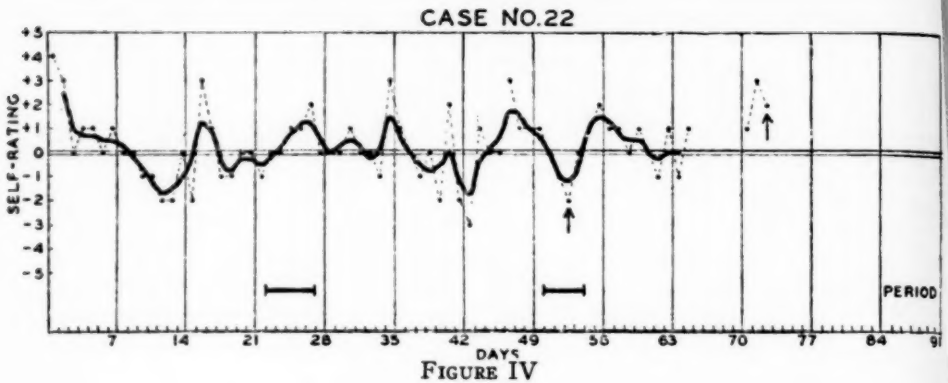


FIGURE IV

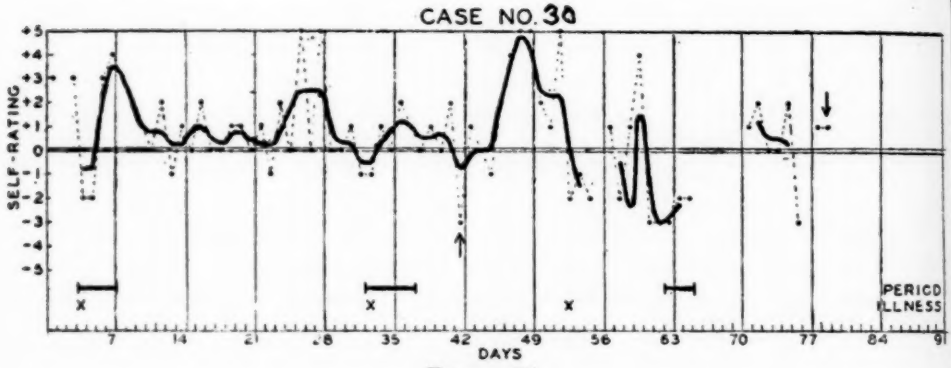


FIGURE V

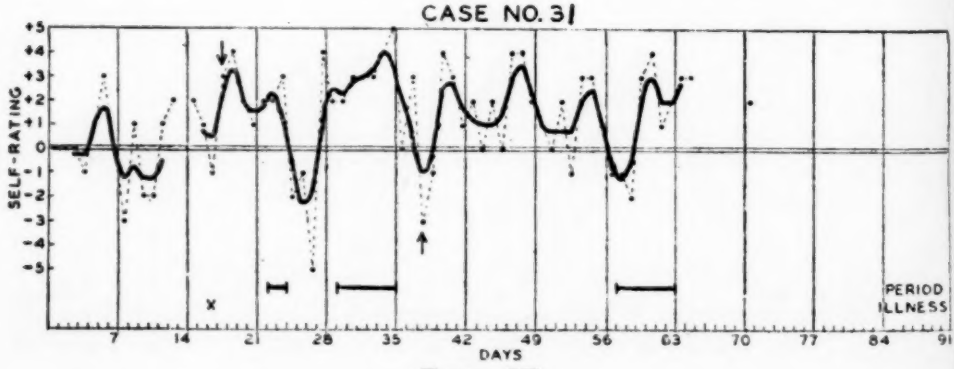


FIGURE VI

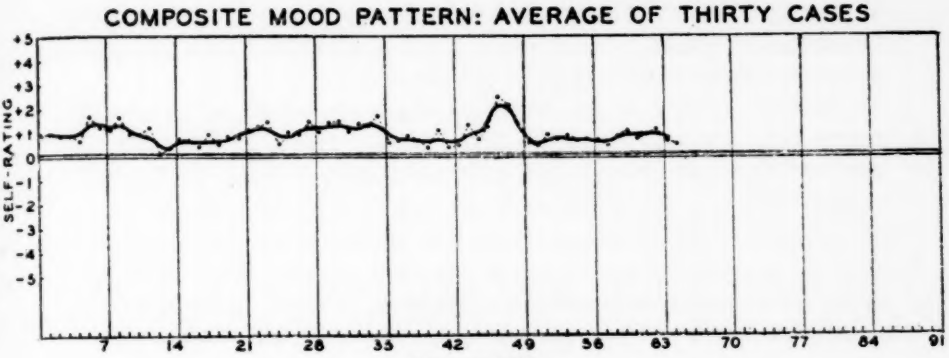


FIGURE VII

present in some patterns, e.g., Case 21. In others it is suggested only by means of the smoothed two-point running average. This is represented by the heavy line which is superimposed upon the original observations represented by dots connected by broken lines.[1]

For many of our subjects a period of about fifteen days between depressions appears to occur more often than chance alone would permit.[1a] Many subjects fail to show changes approximating rhythm in the time within the limits of this study. It does not seem, however, justifiable to conclude that rhythm is not operative for such persons.

Hersey (8) followed twenty male subjects for a much longer period of time than was possible in the present study. He reports "rhythms" for every subject, which vary from subject to subject in length of span from depression to depression. Our own data are insufficient to establish whether or not rhythm is the rule.

A more constant feature of the temporal patterns is the longer persistence of the euphoric as compared with the depressed moods, resulting in a noticeable predominance of "m"-shaped graphs. This phenomenon may be moderately and negatively correlated with age,

[1] The points for the running average, for days b and c respectively, are obtained by the formula:

$$\frac{\frac{a+b}{2}+\frac{b+c}{2}}{2} \text{ and } \frac{\frac{b+c}{2}+\frac{c+d}{2}}{2}, \text{ etc.}$$

This method insures well-weighted points and yet synchronizes extreme points on the average line with observed extremes.

[1a] This periodicity was roughly demonstrated by superimposing all first definite depressions in the thirty graphs. The mean for this synchronized depression was then calculated and likewise the means for following days. This yielded a composite graph which showed a dip at fourteen to fifteen days beyond the point of synchronization. This point was definitely lower than any point on the unsynchronized graph (Fig. VII).

as is suggested by Hersey's data and by the popular notion of the "resiliency of youth." There is, of course, no evidence on this point within the narrow age range of the present study.

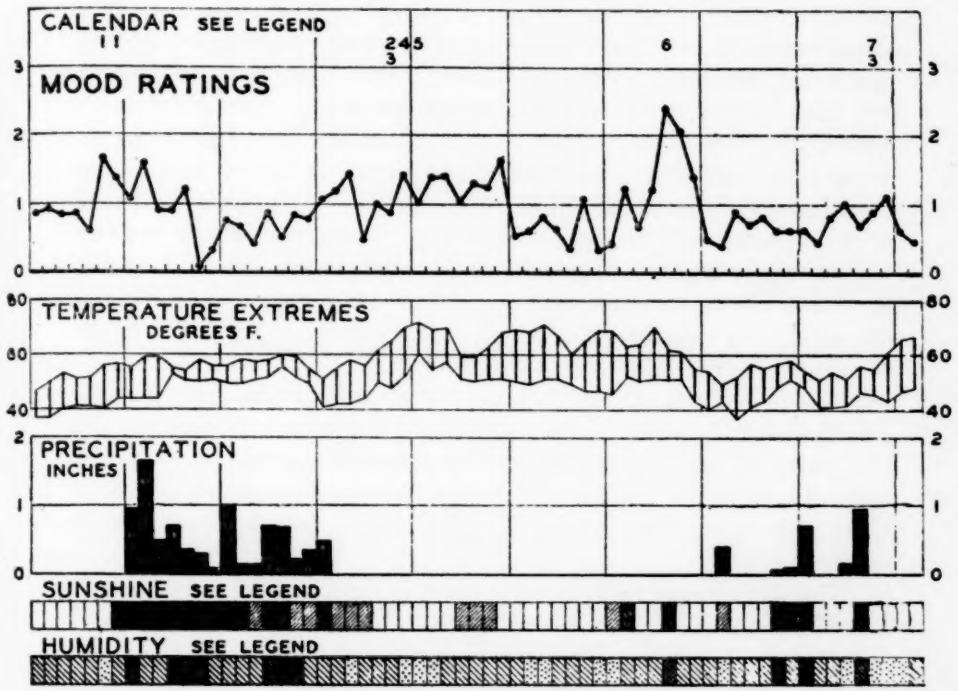
A third feature of the mood graphs is their lack of synchronization with one another. Whatever factors are responsible for the temporal location of rises and falls are largely peculiar to the individual. Nor are these factors in any way obviously connected with the menses. The menstrual period is indicated by the black bars labeled "period," which appear below each graph. Very little evidence for the existence of the "menstrual depression" is forthcoming from these data. The mean rating for the three days before menstruation begins shows an average mood rating of + .60 for the group. The mean rating of the first three days after menstruation begins shows an average of + .50 for the group. These values are very little different from the aveage for all means, which is + .92.; all fall within the 0 to + 1 interval, which is the psychologically neutral area. Other types of illnesses, on the contrary, show a drop in mood levels, the average rating at such times being — .74.

The Composite Mood Graph

The rhythmic appearance of many individual graphs, combined with the lack of apparent synchronization of the rhythm from person to person, suggests that a large factor in mood causation lies in individual physiology. However, many investigators have contended strongly that generalized environmental factors, such as aspects of weather or the time of the week or year, are important.

A graph in which ratings of a group of individuals are averaged is not, strictly speaking, an "average mood" graph at all, since presumably the physiological individualities tend to cancel each other. Such a graph, however, does have the advantage of revealing effects of generalized environmental factors, since it shows what would remain after the cancellation of the unsynchronized individual factors. We have therefore presented in Figure VII a graph representing the composite pattern of our group. A point on this graph represents the average of all subjects in mood rating for the day. The majority of the points lies within the 0 to + 1 interval. Only in three sections does the group indicate any particular deviation above this area. Nor can one read any great amount of mood change into these slight fluctuations. The average mean is .91 and the average sigma 1.86. Even response to the formal dance shows less rise than one average sigma. Mood changes have largely disappeared.

Weather Variables Compared with Enlarged Composite Mood Rating



LEGEND

CALENDAR

1. INITIATION
2. BIG "C" CIRCUS
3. END OF EXAMINATIONS
4. SUNDAY NIGHT SUPPER
5. INSTALLATION OF OFFICERS
6. FORMAL DANCE
7. FOUNDERS' DAY LUNCHEON

SUNSHINE
PER CENT OF POSSIBLE HOURS

SUNNY	75-100
MODERATE	50-74
CLOUDY	0-49

HUMIDITY
PER CENT

DRY	0-49
MODERATE	50-74
HUMID	75-100

FIGURE VIII

Figure VIII, in which the composite pattern is reproduced on an enlarged scale together with notations of social events and graphs of weather factors, indicates that there is very little evidence that any weather factor is correlated with mood. Only the most glamorous social occasion affects the graph in any probably significant fashion. There are no regularly blue Mondays, February is not more depressing than March or April, and sunshine has a negligible effect. The writer is aware that much elaborate work has been published on weather as affecting biochemistry and mood (12) but has not yet seen any experimental work in which mood is treated as a variable which would negate the present findings. It is true that our subjects are unusually healthy, busy individuals upon whom such factors may have less effect than upon other populations.

B. The Behavior Correlates of Mood

After the subjects had become accustomed to the daily routine of supplying the self-ratings of euphoria-depression as described in the preceding section, individual experimental sessions were begun. All data for both the subjective and behavioral analysis of mood were gathered at these sessions. Each session lasted approximately ninety minutes.

The experimenter chose the subject in each case on a dual basis: (1) The subject's euphoria-depression rating for the first day must be above or below the 0 to + 1 interval on the self-rating scale.[2] (2) The time must be convenient for the subject. It was considered important that this be stressed so that subjects should not be penalized for extreme self-ratings.

Two experimental sessions were taken with each subject: one on a day when the self-rating indicated a pleasant mood, and one on a day when the subject gave a depressed mood rating. The time interval between sessions was not limited. The experimenter strove, of course, to obtain samplings with as wide a spread as possible between ratings. Actually the minimum difference between any two sessions taken was approximately two sigmas. All experimental sessions were held at the same period of the day, in the evening. The examination with euphoric mood preceded that with depressed mood in the case of fourteen subjects. The reverse was the order in the case of the remaining fourteen subjects.[3]

[2] One or two experimental periods were held at ratings of -1 in cases where the opposing rating was extreme.

[3] Two of the thirty subjects who provided daily mood ratings failed to complete the experimental sessions upon which analysis is based.

The attempt to construct a battery of tests which would demonstrate reliable behavioral differences between pleasant and unpleasant moods in well-integrated, highly socialized, and highly intelligent normal subjects presents many difficulties. Two points seem especially important. (1) The measure must be relatively inaccessible to the effects of suggestion. This is especially desirable, since suggestion is not ruled out in the later analysis based on subjective rating of mood components. (2) The measures used must be somewhat indirect, since the general consensus of opinion is that in any directly presented task the normal subject temporarily compensates for any decrement due to such ephemeral factors as mood by a greater effort.

For this section then, we wished to collect indirect measures of such behavior trends as differentiate depressed from elated patients. If the behavior of these normals can be shown to vary reliably with their mood changes in the direction expected from psychiatric findings with abnormals, it seems reasonable to consider that a continuum between normal and abnormal moods has been demonstrated.

Experiment I. Spontaneous Vocalizations

The measure of vocalization was considered important, since some degree of retardation of speech, with the extreme as mutism, is a clinically common correlate of depressed states. For example, the English factorialists, working with psychotics in investigations of their "f" (flow or fluency) factor—which they judge to be central in mood—have concluded that "F" measures the amount and speed in talking rather than the speed of simple association. Manic response to an association test may be slower than usual, while "f" scores are consistently high (12a).

The present study is based upon a covert count of spontaneous remarks made by the subjects during the course of each experimental session. Since the period was well filled with tasks requiring the subject's attention and only brief unitary remarks were possible, this measure proved relatively easy to score correctly. A "remark" was taken to be a statement or question, but exclamations were excluded.

Subjects were found to average about 61 per cent of their remarks while in a cheerful mood.[4] Statistical treatment yields a critical ratio of 4.96, indicating a probability of much less than one

[4] See Table I, p. 96.

in a hundred that this difference could have occurred by chance.[5] It therefore seems probable that for similar populations the talkative-silent tendency may be considered a mood variable.

Experiment II. Speed of Associations

The literature and clinical findings from the abnormal field suggest that speed of associations should be differential for high and low moods. The flight of ideas in the manic contrasts markedly with the retardation of thought in schizophrenia and with the blocking in circular depressions. Studman (13) has found with normals that the easy touch off of associations is central for high score on the "f" factor. Fisher and Marrow (5) confirm this finding.

In our study, word associations were secured in response to stimulus words given orally by the experimenter. All responses given within one minute after each stimulus word were taken down as the subject pronounced them. Six stimulus words, chosen from a preliminary study, were used (9). These were: house, black, time, man, ambition, surf. The same list was used for both sessions. Scores were based on the number of responses made during the six-minute euphoric sampling as against the number made by the same person during the six-minute depressed sampling. Since the experimenter wished to make every effort not to suggest the rate of response, the stop watch was concealed. The precise directions used are given below.

DIRECTIONS: "As you know, we are studying your day-to-day feelings and emotional attitudes; memories and word associations help us to do this. I want you to close your eyes and allow your mind to be dominated by your feelings. Now I shall say a word to you, and when I do so I want you to say the first word that my word makes you think of (give single words, not phrases). Your word will bring another thought to your mind, and that word will bring another, so that you can give me a chain of ideas that goes on until I have all I need. Now relax, let your feelings dominate, and when I suggest a word say the *first word* that comes. As your mind runs on, don't try to choose *suitable words;* it's the *first-rising* thoughts we want."

Results of the experiment yielded a difference of less than reliable magnitude (critical ratio 1.71). Fifty-two per cent of associations occurred during euphoria. In the author's opinion the aim of measuring unforced associative speed was not satisfactorily achieved. Clinical notations (made throughout the study by the

[5] Following Fisher's Tables for an "n" of 28, we accepted any critical ratio in excess of 2.5 throughout this study as significant.

experimenter) indicate that in many cases effort was put forth by the subject to speed up slow associations. That is, the subject conceived of the directions to "give associations" as a *task*. The measure became too direct, and the mood effects were attenuated.

Studman, working with psychotics, has used a similar measure by urging "utmost speed" on her subjects (13).

Experiment III. Emotional Tone of Associations

It may seem almost a truism to state that in an euphoric mood the affect of the associative content should be definitely more pleasant than in depressed moods. Bleuler's (2) concept of catethymia is based on this. Indeed, in our earlier study (9), it was somewhat naïvely assumed that a validation of the criterion of mood was here available. Reflection, however, makes it quite apparent that the matter is not so simple. The very fact of present euphoria, for example, might conceivably bring about more annoyed than pleasant reaction to a short sample of stimulus words, provided, for example, that the subject was preoccupied with an euphoric estimate of his abilities in relation to the small recognition accorded them. Fankhäuser (4), too, has pointed out that elation may occur with negative affect—as in disgust or discontent. The obverse "rose spectacles" reaction is also a possibility in which a genuinely depressed subject finds pleasant affect to accompany thought trends dealing with an irretrievable past or an unattainable future.

In the present study, at the close of the six-minute samplings of associations, the experimenter repeated the subject's responses slowly one by one in the order given, and asked the subject in each case to indicate whether the accompanying affect was pleasant or unpleasant. If the affect seemed neutral at first, the subject was instructed to repeat the association mentally in the original connection several times, until it took on definite emotional tone. Precedent for this procedure is found in Washburn's studies (14). Scores were based on a comparison of pleasant and unpleasant associations.

DIRECTIONS: "Now let us go back and remember what feelings you had as you thought those thoughts. I will repeat them to you one by one, and as I do you say 'pleasant' or 'unpleasant,' depending on whether the feeling you had when it came to you was more pleasing or more unpleasant or annoying. If a word seemed neutral at first, repeat it to yourself until it has a definite pleasant or unpleasant connotation. Please give the real feeling you had when you said the word, not what you think you should have, nor necessarily the feeling you would usually have to that word."

Our results indicated that 52 per cent of pleasant associations occurred in euphoric mood. This amounts to a trend of difference only, but the direction is the same as shown by Bousfield's (3) data based upon larger numbers.

Experiment IV. Regression in Thought Trends

This was suggested by an a priori assessment of the nature of depressed mood. Euphoria-depression appears to the author to be psychologically the expansion-retraction trends of the ego.

At the close of the experiment on emotional tone of the associative responses (Experiment III), the experimenter slowly repeated each response once more, and the subject indicated whether or not the response referred to an event in "childhood." Childhood was arbitrarily defined as "that period before you finished grade school." The score was percentage of "old" associations in the euphoric mood as against the percentage in the depressed mood.

DIRECTIONS: "Now we want a rough estimate of whether the thought that was in your mind when you gave the associations was an *old* memory or whether you were thinking of something rather contemporary or future. I will repeat your responses again, and you say 'old' if the idea you had rose out of experience in childhood or before high school."

Our subjects, when depressed, indicated definitely more associations which returned to childhood. Only 37 per cent of all regressive thoughts occurred during euphoria. The probability of the found difference arising by chance is less than one in a hundred.

Experiment V. Speed of Decision

This experiment was suggested by psychiatric literature and by clinical observation of depressed patients showing retardation of thought process; also by the author's subjective experience that in depressed moods the ability to come to a decision seems retarded. We wanted data to yield evidence on whether or not the subjective feeling of retarded judgment in depression is more than an illusion. Considerable confidence was placed in the freedom of this item from either of the most common dangers to mood findings: (1) the effects of suggestion and (2) the dissipation of differences by compensatory effort on the part of the subject. Our subjects seemed absorbed only with the "rightness" of their judgments. All clinical notations made by the experimenter during the testing indicated that the measure was satisfactorily indirect.

Five pairs of small, weighted boxes were used for this experiment. All boxes were identical in appearance; members of a pair were identical in weight. Mild deception was resorted to in the instructions, as indicated below. Time taken for each decision was recorded. Scores were based on the total time for decision.

DIRECTIONS: "Here are five pairs of weights. The pairs are not quite equally heavy. Take this pair and heft them as you please until you can decide which you think heavier."

Results indicate that decision was reliably retarded under depressed mood as against decision made in euphoria. Only 44 per cent of total time expended in making decisions was so expended during euphoria. There is less than one chance in a hundred that the found difference is due to chance.

Experiment VI. The Expansive-Retractive Tendency in Judgment of Distance

This experiment was suggested by the observed fact that elated and excited psychopaths make gestural responses which are expansive, wide, outgoing, or centrifugal in type, while gestures characterizing very depressed, anxious patients are meager, close-held, and centripetal. Münsterberg (11) reports mood differences on distance judgments in self-experimentation. Since, however, he was aware of his own hypothesis, unconscious suggestion may have been operative. Vernon and Allport (1) have used this measure in their study of congruence within the personality.

We asked the subject to stand up and close her eyes with palms of hands together at the chest. She was also blindfolded as a precaution against inadvertent cues; she was then asked to move her hands apart a distance which she believed to equal one foot. This distance was measured. The experiment was repeated with distances of eighteen inches and of two feet. This process was carried out during both experimental sessions for all subjects.

DIRECTIONS: "I shall blindfold you now, and after that I want you to hold your hands in this position [shows] at various distances from your body as I will tell you. Now hold your palms parallel like this [shows] and place them at varying distances apart as I tell you."

Scores are based on the magnitude of estimates in euphoric mood as against those in depressed mood. The chances are approximately 96 in 100 that the difference is a true difference. Fifty-two per cent of all linear distance indicated was indicated during euphoria. The

mean estimate in euphoric mood is 57.23 inches, in depressed mood 53.45 inches. The mean for the group is an overestimate in euphoric mood and an underestimate in depressed mood, since total distance was fifty-four inches. The amount of over- or under-estimate is probably a function of the personality make-up of the group involved.

Experiment VII. Psycho-Motor Facilitation

VII a. Expansion-Retraction in Digit Area. European graphologists have stated that mood factors interfere with consistency of handwriting characteristics (10). Flow and scope of learned neuromuscular responses would seem a priori to be related to the surgency concept of the factorialists. Since motor retardation is so common a clinical accompaniment of psychopathic depression, the experiment on size of figures promised well.

We used a large sheet of unlined brown wrapping paper, eighteen inches by twenty-one inches, which we gave to the subjects with the instructions: "Write out numbers going backward from 100 [demonstrate] until I have as many as I need. I'll say when to stop." The backward order was used in the hope of masking the true purpose of the measure. The unusual paper was employed in both writing measures to avoid the habit pattern of size evoked by standard foolscap. Score was based on the area covered by the figures produced in one minute. The area was boxed length times the average width of columns (measured to the nearest one-fourth inch).

The figures produced in euphoric moods proved definitely more expanded than those produced in depressed moods, the probability being less than .01 that the difference was due to chance. The critical ratio is 3.25.[6]

VII b. Expansion-Retraction in Written Area.—It was expected that written samples might show less definite mood effects than did Experiment VII a (in spite of the use of non-standard size paper and indirect approach), because habit patterns set up by copy and ruled paper in childhood are perhaps more rigidly set by extensive practice than is the case with size of figures.

This part of the experiment was announced as a memory test.

[6] This finding is of added interest in view of the fact that when the test is scored for efficiency (as measured by the actual number of digits produced) only a trend of difference is demonstrable. The critical ratio is 1.44. This last finding was anticipated by the experimenter, since it was obvious during the experimental sessions that in spite of leisurely instruction and a concealed stop watch, many of our subjects were interpreting the instructions as a speed task.

DIRECTIONS: "See how much you remember of the ideas expressed in the paragraph I now read to you." A standard paragraph of interesting content (from Benedict's *Patterns of Culture*) was read to the subject. She was then presented with the reverse side of the sheet of wrapping paper used in the previous experiment and was asked to write down as many of the ideas in the selection read as she could remember. No time limit was used. Two roughly comparable standard paragraphs were used in the study. Half of the subjects were presented with paragraph one in euphoric mood, and the other half used paragraph two. The order was reversed in the depressed moods. Scores were based on the average length of line, multiplied by the linear vertical distance from the top of small letters in the top line to the bottom of small letters in the lowest line.

The difference is in the expected direction but is not entirely reliable. There are approximately five chances in one hundred that the found differences might occur by chance.[7]

VII c. Expansion-Retraction in Gesture.—Two items were included under this heading. The subjects were asked to draw on a large wall-board three circles and later three squares.

DIRECTIONS: "I would like you to draw a circle on the blackboard—— Make another, please—— and one last one. Now pretend that you have erased these and draw a square—— one more—— and a last one."

The scores were areas of "squares," and horizontal-times-perpendicular diameters of "circles." Squares drawn in euphoric mood were scored against squares drawn in depression, and similarly with circles. Results were in the expected direction but showed small critical ratios: for squares, 1.92; for circles, less than 1.

Experiment VIII. Perceptual Objectivity; Rate of Perceptual Fluctuation with Ambiguous Figures

This experiment was included because of the interesting work of Guilford and Hunt (7), who determined the rates of fluctuation of an outline cube for manic-depressive and schizophrenic patients. They wished to test McDougall's theory that introversion and extroversion are manifested respectively in high and low rates. Schizophrenic patients showed an average rate about four times that of the manic-depressive group.

[7] When these data were scored simply for the flow of words, the difference in favor of the euphoric mood is not significant (critical ratio 1.15). This is in line with the findings for surgency-bloc obtained in Experiment II (Associative Flow).

In the present experiment familiar ambiguous figures were used: (1) the reversible-stairs figure, (2) the piled-cubes figure. Four one-minute samples of rate of fluctuation were given at both experimental sessions. The figures were presented in 1, 2, 2, 1 order. Any subject who was not familiar with the phenomenon of perceptual fluctuation was allowed to experience it before beginning the experiment. Score was based upon number of reversals in depressed mood as against the number in euphoric mood.

DIRECTIONS: "Here are two ordinary ambiguous figures, which appear to change when one continues to look at them. Please watch them and whenever they change over you say 'change'. Watch until I tell you to stop."

The mean number of fluctuations in euphoric mood was forty-eight as against fifty-five in depression. Forty-six per cent of all fluctuations were found in euphoric mood. The critical ratio is 2.24, indicating approximately ninety-seven chances in one hundred that the fluctuations in euphoria are less than in depression. It is noteworthy that depression yields the more rapid scores.

This finding, taken in conjunction with that of Hunt and Guilford (7), is an interesting bit of evidence for our hypothesis. It also falls in line with the current work of George (6), who finds that extreme introverts show more rapid fluctuations than do extreme extroverts and that the rate is affected by drugs.

Experiment IX. Sleep Reports

Manic and elated patients are known to sleep less and more brokenly than apathetic or depressed patients. Hersey (8) reports a difference in the same direction with normals under euphoric and depressed moods.

Reports were requested at each experimental session on the nature of the sleep of the past night. The number of hours of sleep, whether it was broken or unbroken, and the presence of dreams were recorded. Euphoric subjects reported a total of 210 hours of sleep on the nights recorded, depressed subjects a total of 219 hours. Euphoric subjects reported 17 nights of sound sleep and 11 nights of restless or broken sleep. Depressed subjects reported 22 nights of sound sleep and 6 nights of restless or broken sleep. Euphoric subjects reported 12 nights when dreams were present, while depressed subjects reported 8 nights when dreams were present.

The sleep conditions of our particular group of subjects were so unusually subject to environmental forcing that their comparability

with the sleep needs or desires of the organism is doubtful. The difference is unreliable, but the direction is in line with the hypothesis.

C. Treatment of Behavior Analysis Scores

A proportional method was adopted for the comparison of behavior scores in euphoric as against depressed moods. The rationale of the formula evolved was as follows: if there is no difference in the performance, 50 per cent of all responses will occur in euphoric mood and fifty per cent in depressed mood (except for differences due to chance). Therefore, if "p" represents the proportion of responses occuring in euphoric mood and "q" represents the proportion of responses occuring in depressed mood, then from the critical ratio

$$\frac{Mp - .50}{\frac{\sigma p}{\sqrt{28}}}$$

the probability of the euphoric mean being greater than .50 may be determined. If the scores in depressed mood are greater than in euphoric mood, the critical ratio

$$\frac{.50 - Mq}{\frac{\sigma q}{\sqrt{n}}}$$

will determine the probability of the depressed mean being less than .50. This formula was evolved,[8] because the nature of the data was such that a certain amount of distortion of the facts appeared to result from the application of the more usual method for computing the reliability of a difference. In these data "euphoric" and "depressed" are relative terms—that is, they refer in each case to the norm of the particular individual who is rating herself. The behavior scores, on the contrary, if handled as raw scores, are treated as absolute values having reference to a general norm. Very wide personality differences are involved, and thus a measure such as the mean becomes unduly unstable. An example will make the situation clearer. In the experiment where the score is number of spontaneous unnecessary remarks made during an interview, subject "A" makes ten such remarks during the euphoric interview and five during the depressed interview. The difference, as crude score, is five. Subject "B" makes forty remarks during euphoric interview and twenty remarks during the depressed interview. The difference, as crude score, is twenty; and subject "B" contributed undue weight toward the determination of the mean difference. Yet there is no difference between "A" and "B," when one considers them in terms of the question—"Does a person become more talkative when he is in an euphoric mood than when in a depressed mood?"—which is the question this experiment attempts to answer.

[8] Acknowledgment is due to Professor Robert Tryon, who suggested this formula when the author expressed reasons for dissatisfaction with the usual technique.

The proportional method gives more weight than does the use of crude scores to differences which are proportionally large but absolutely small and, for this reason, seems better adapted to these data.

TABLE I

BEHAVIOR DIFFERENCES BETWEEN EUPHORIC AND DEPRESSED MOODS

Experiment	Mp*	σ p	σ Mp	$\frac{\text{Mp-50}}{\sigma \text{ Mp}}$
I. Vocalizations	.614	.118	.023	4.96
II. Flow of Associations	.524	.074	.014	1.71
III. Affect of Associations	.520	.068	.013	1.54
IV. Thought Regressions	.371(Mq)*	.205(σ q)	.039	3.31 $\frac{\text{50-Mq}}{\text{mq}}$
V. Decision speed	.448	.091	.018	—2.89
VI. Expansion in Judged Dist.	.518	.040	.008	2.25
VII. Psycho-Motor Facilitation				
(a. digit area)	.552	.084	.016	3.25
(b. writ. area)	.552	.133	.026	2.00
(c. draw. sqs.)	.548	.125	.025	1.92
VIII. Perceptual Fluctuation	.462	.086	.017	—2.24

* (p = proportion of total reactions occurring in euphoric mood $\frac{E}{E+D}$.

(q = proportion of total reactions occuring in depressed mood $\frac{D}{E+D}$.

The measures taken on the above functions (done once in euphoric and once in depressed moods) were as follows:

I. Number of spontaneous remarks made during interview.
II. Number of associations given in a six-minute period.
III. Number of associations in above periods which were pleasant or unpleasant.
IV. Ratio of total associations to childhood memories in above periods.
V. Time in seconds taken to make five decisions in case of an unsolvable problem.
VI. Magnitude of judgments when three estimates of distances between hands is requested (blindfolded).
VII. Figures, writings, and drawings measured for area.
VIII. Number of fluctuations in ambiguous figures in four trials of one minute each.

D. SUMMARY OF EFFECTS OF MOOD ON BEHAVIOR

Under the conditions of this study certain kinds of behavior seem very surely affected by the mood of the subject. The subjects certainly made more spontaneous and unnecessary conversation in euphoric than in depressed mood. Depression resulted in a very definite regression to childhood events, in thought and memory, compared to the amount of this type of association in euphoria. Our subjects found it possible to make decisions much more readily when in pleasant moods. Apparently the sense of indecisiveness and mental futility which often accompanies depression is not entirely illusory. Our subjects also showed a definite change in such psychomotor functions as writing figures; in euphoric moods they made larger, more expansive movements in such work than when depressed. All critical ratios supporting these changes indicate a probability much greater than ninety-nine chances in one hundred that the difference is a true difference (Table I).

Other behavior changes occurred to an extent which indicates that the found differences are probably genuine; in these cases the chances being ninety-five or ninety-six in one hundred that the difference is reliable. Depression probably causes retractive trends within the individual which lead him to judge distances to be greater than he judges them when in pleasant moods. Our subjects, when making a gesture which was supposed to mark off a given length, underestimated (on the average) in depression and overestimated in euphoria. Apparently script is inclined to be cramped and smaller when one is depressed. The change in size is less, however, than that found for figures. Perceptual fluctation is the final behavior variable demonstrated to vary with mood to an extent which indicates reasonable reliability. This phenomenon of change in ambiguous figures may be variously thought of as dependent upon the degree of central tension or as an ability to objectify experience. Its nature is not definitely known, but the changes in perception tended to proceed at a faster rate when our subjects were depressed.

Other measures, such as the tendency in depression to show a slower speed of association, to have less pleasant associations, and to make smaller gestures, as in drawing a square, yielded differences of less than significant magnitude. The report on sleep suggests that it may be more profound in depression. Interpretation of behavioral data will be undertaken in the section on the nature of mood (Part II).

References

1. Allport, Gordon and Vernon, P. E., *Studies in Expressive Movement,* New York, 1933.
2. Bleuler, E., *Lehrbuch der Psychiatrie,* Berlin, 1920.
3. Bousfield, W. A., Qualitative correlates in feeling tone, *A.P.A. Program,* Sept., 1936.
4. Fankhauser, E., Gefühl, Affekt und Stimmung, *Zsch. f. d. Ges. Neur. u. Psychiat.* 1931, **132**, 321 ff.
5. Fisher, V. E. and Marrow, A. J., Experimental study of moods, *Char. & Person.* 1934, **2**, 201-208.
6. George R. Wilfred, Significance of the fluctations experienced in observing ambiguous figures and in binocular rivalry, *J. Gen. Psychol.,* 1936, **15**, 31-35.
7. Hunt, J. McV. and Guilford, J. P., Fluctuations in an ambiguous figure in dementia praecox and in manic depressive patients, *J. Abn. & Soc. Psychol.* 1932-33, **27**, 443-452.
8. Hersey, Rexford B., *Workers' Emotions in the Shop and Home,* Philadelphia, 1932.

9. JOHNSON, W. B., Effect of mood on personality traits as measured by Bernreuter, *J. Soc. Psychol.* 1934, **5**, 515-522.

10. LEIBL, M., Der Stimmung Ausdruck in der Handschrift, *Zbl. Graphol.*, 1935, **5**, 237-252.

11. MÜNSTERBERG, H., Die Psychophysische Grundlage der Gefühle, *Proc. II Internat. Cong. of Exper. Psychol.*, pp. 132-135. London, 1892.

12. PETERSON, W. F., *The patient and the weather*, Vols. I, II, III, IV, Ann Arbor, Mich., 1937.

12a. STEPHENSON, W. et al., Spearman factors and psychiatry, *Brit. J. Med. Psychol.*, 1934, **14**, 101-135.

13. STUDMAN, L. G., The factor theory in the field of personality, *Char. & Person.*, 1935, **4**, 34-35.

14. WASHBURN, M. F., A comparison of directed and free recalls of pleasant and unpleasant experiences as tests of cheerful and depressed temperaments, *Amer. J. Psychol.*, 1926, **37**, 278-280.

THE STANFORD-BINET TEST AS A PSYCHOMETRIC METHOD

P. E. VERNON

Jordanhill Training Center, Glasgow

I. The Psychometric Approach

The appearance of the revised Stanford-Binet Test[1] provides a good opportunity for ventilating the controversy between those who criticize this test as an antiquated and unreliable instrument which should give place to more modern and scientific techniques, and those who on the other hand persist in regarding it as the best psychological measuring instrument at their disposal. In a recent discussion,[2] Dr. R. B. Cattell put forward the former view, but he was strongly opposed by many of the experienced Child Guidance Clinic psychologists who were present. Although I must not claim that my views are representative of those held by Clinic psychologists in general, yet I hope, in this article, to place our opposition upon a more systematic footing.

The controversy would appear to result from a still more fundamental disagreement between those who favor what may be termed the *psychometric* and the *clinical* approaches to the study of personability. As I have indicated elsewhere,[3] both approaches are legitimate and valuable, and they should reinforce rather than exclude one another. But the psychometric approach by itself is unable to

[1] L. M. Terman and M. A. Merrill, *Measuring Intelligence* (Cambridge, Mass., and London, 1937).

[2] *Report of the Inter-Clinic Conference* (London, 1935), pp. 24-29.

[3] P. E. Vernon, Can the total personality be studied objectively? *Charact. & Person.*, 1935, **4**, 1-10. The significance of the Rorschach test, *Brit. J. Med. Psychol.*, 1935, **15**, 199-217.

I have been so frequently credited with supporting wholeheartedly the clinical or intuitionist standpoint and opposing the psychometric standpoint that I should like to reiterate what was stated in these and other publications, namely, that my desire is to reconcile and combine them so as to bring about a broader (but not less scientific) study of personality. I do in fact greatly respect Dr. Cattell's viewpoint, and accept many of his criticisms of the Binet testing procedure; and I am grateful to him for writing, at my request, the article which appears in this issue.

satisfy the requirements of psychologists engaged in clinical work. I must discuss this point first, and will then hope to show that the Binet Test is a very successful compromise between the two approaches.

The psychometric approach is characterized by two essential features: (1) rigid standardization of the testing procedure, together with empirical proof of reliability and validity, so as to eliminate the uncertainties and the bias inherent in more subjective methods of assessing human beings; (2) the conception of traits or abilities as discrete variables, whose variations among different individuals are purely quantitative. No reliance is placed on subjective ideas of the nature of human traits; their existence as true "vectors" of the mind should be established by the methods of factorial analysis.

For example, the common notion of practical ability cannot be accepted merely because in everyday life we regard some people as good, others as bad, at practical tasks. It is necessary first to obtain accurate measurements of the performances of a group of people at a variety of practical tasks, then to determine by correlating their scores whether this practicality is general (whether those who are good at one task are also good at others); and thirdly, to examine whether such generality or overlap is not due merely to the fact that each task involves a certain amount of general ability or "g" factor.

The writings of Kelley,[4] Cattell, and others who support the conception of unitary traits, indicate that they hope ultimately to isolate all the chief factors underlying the multitudinous traits and abilities which distinguish human beings, and to be able to measure them separately. Such factors might be compared to the chemical elements, and each individual regarded, like a chemical compound, as a combination of elements in certain proportions. The tests which we employ at present usually involve several elements (in addition to specific factors) and scholastic or vocational success likewise depends on a certain pattern or formula of factors. Thus by applying a comprehensive battery of tests whose factor content had been analyzed, we should be able to calculate the individual's standing on each of these primary traits; and having already determined the factor content of various lines of work, we should be able to deduce objectively what are the lines for which his factor make-up

[4] T. L. Kelley, *Crossroads in the Mind of Man* (Stanford, Calif., 1935); *Essential Traits of Mental Life* (Cambridge, Mass., 1935). Cf. also C. L. Hull, *Aptitude Testing* (Yonkers, N. Y., 1928). D. G. Patterson et al., *Minnesota Mechanical Ability Tests* (Minneapolis, 1930).

fits him. By such means vocational guidance and child guidance might at last become truly scientific.

Possibly this bald statement of the aims of contemporary psychometrists is unintentionally exaggerated; few writers may subscribe to so mathematical a conception of human nature and mental measurement. Nevertheless, it would appear to be the logical outcome both of the criticisms of the unscientific nature of Binet testing, and of much of the recent work on the factorial analysis of abilities.

II. The Clinical Viewpoint

The clinical psychologist cannot agree with either of the psychometric principles enunciated above. He doubts the efficacy of standardizing the objective conditions of testing, since it may fail to standardize the "subjective situation," i.e., the meaning of the situation to the person tested; and it is this meaning which will determine the test response. He knows that in order to influence or understand different individuals he has to treat each one differently, and so concludes that a group test or other objective psychometric method, where all individuals are treated alike, cannot reveal much that is of value.

Neither the clinician, nor in all probability the layman, makes much use of the conception of quantitative variables. In everyday life we do indeed quite often apply it, e.g., to school work, athletic contests, income, etc.; but in our direct dealings with one another, qualitative distinctions appear to predominate, and single vectors of mind are seldom considered in isolation from the complex structure of the personality. Thus the separate treatment of g, intelligence, or other abilities, as purely cognitive variables seems to the clinician to involve a process of abstraction which is untrue to the psychological facts. Only when such measures are considered synthetically with other information about the personality can they tell him anything about the individual which will be of assistance in diagnosis and treatment.

Our judgments of, and our reactions to others are initially based on their features, gestures, etc., or on observations of their specific acts; but we interpret these, partly by logical (or illogical) inference, partly by "intuition," as significant manifestations not merely of discrete traits but of the total personality structure. We may class any one such act under a certain trait-name, or consider it as indicative of a certain ability; but primarily it is a sample of the

whole person as we see him. The greater the number and the variety of samples that we are able to observe, the better we feel that we know him. That such inferences and intuitions are often faulty is obvious. Yet they continue to subserve the majority of our practical adjustments to one another, and play an overwhelming part in the clinical approach.

III. Discussion of Evidence

The objections outlined above are, we must admit, derived largely from the personal experience of clinical psychologists. Yet they have also been put forward by experimentalists. Bartlett, for instance, has criticized the dogma of standardized conditions;[5] Peters[6] and M. Bleuler[7] have argued against the separation of intelligence from the rest of personality. And many lines of supporting evidence might be adduced.

Space precludes a full discussion here of factorial analysis. I have given reasons elsewhere[8] for concluding that it is of the utmost value for classifying mental tests, both cognitive and orectic, and for generalizing the results of correlational studies of traits and abilities; but that it cannot justifiably claim to reveal the underlying psychological structure or real elements of personality. Such a conclusion agrees with the view which, according to Holzinger,[9] "is now held by all members of the Unitary Traits Committee." Thus the description of people in terms of factors is for certain purposes both possible and useful. But the clinician is not violating any scientific obligation because he finds such description inadequate for his work, and relies on an entirely non-mathematical conception of the mind.

I must treat rather more fully the question of the legitimacy of isolating and measuring separately the different vectors of mind. Is it true that each response to an intelligence test reveals not merely intellectual capacity but also reflects the total personality structure, and, conversely, that the subjective state of mind, the moods and emotional traits of the testee greatly affect the adequacy of his test responses?

[5] F. C. Bartlett, *Remembering* (Cambridge, 1932).

[6] W. Peters, Das Intelligenzproblem und die Intelligenzforschung, *Zsch. f. Psychol.*, 1922, **89**, 1-37.

[7] M. Bleuler, The delimitation of influences of environment and heredity on mental disposition, *Charact. & Person.*, 1933, **1**, 280-300.

[8] P. E. Vernon, *A Survey of Attitude Tests, Rating Scales and Personality Questionnaires* (to be published shortly).

[9] K. J. Holzinger, Recent research on unitary mental traits, *Charact. & Person.*, 1936, **4**, 335-343.

Spearman[10] has discussed this point, and has quoted a number of experiments which fail to show any appreciable effect of orectic processes on group test results, and therefore on g. He concludes that "Insufficiency of will-power, nervous temperament, and susceptibility to complexes do handicap people in the prolonged, familiar, and monotonous occupations of ordinary life, but have no effect in the brief, novel, and interesting performances constituting the tests." The experiments do not appear, however, to be very conclusive; for adequately controlled investigations of human motivation are not easy to carry out.* Moreover, many lines of evidence do suggest the existence of an intimate connection between ability and orexis.

There is first a small but indubitable positive correlation between intelligence and emotional stability, qualities of character and maturity of interests. Terman's studies of gifted children point in this direction, as also do the many investigations of the low I.Q.'s of delinquents. Again maladjusted children in general, who attend the ordinary child-guidance clinic, are somewhat below average.[11] The marked tendency to emotional deviations in mental defectives is well known; indeed, Earl[12] is led to conclude, from his study of cases of extremely low intelligence, that "intelligence . . . is built upon a basis of affective-instinctive factors."

This general relationship between emotional and intellectual characteristics is less striking than the specific subnormality of the maladjusted on particular tests. I have recently shown that clinic children and adults score much lower on the Seguin-Goddard Formboard and the Healy Picture Completion II than on the Binet Test, and have adduced other similar experimental results.[13] Porteus claims that his maze test depends on social adjustment in addition to intelligence. And Jastak[14] has found that abnormal children, who were equal to a control group of normal children on the Vo-

* This objection to the experiments seems too indefinite to be very helpful. But possibly the author has not taken into account that the statement quoted by him refers only to certain particular influences ("nervous temperaments," etc.). *Other* orectic troubles, such as many kinds of insanity, or even of minor mental disorder, can doubtless make testing extremely difficult, or indeed illusory.—*Editor.*

[10] C. Spearman, *The Abilities of Man* (London, 1927).

[11] Cf. P. E. Vernon, A study of the norms and the validity of certain mental tests at a child guidance clinic, *Brit. J. Educ. Psychol.*, 1937, **7**, 72-88, 115-137.

[12] C. J. C. Earl, The affective-instinctive psychology of imbecile children, *Brit. J. Med. Psychol.*, 1936, **15**, 266-278.

[13] *Op. cit.*

[14] J. Jastak, *Variability of Psychometric Performances in Mental Diagnosis* (New York, 1934).

cabulary test, were much inferior to them on a maze test, on Cube Construction, and on certain of the Binet items—Reading and Recall, Digits Backward, and Memory for Designs.

Factorial studies of abilities have also indicated the influence of orexis. In Alexander's investigation,[15] certain group factors were discovered and presumed to be affective in nature, which played a large part in scholastic and vocational achievements, and which also entered to a smaller extent into intelligence test correlations. Hamley[16] has described factorial studies among schoolchildren, where an attempt was made to estimate the interest or motivation among the subjects. It was found that when this was partialled out, the correlations and the factor patterns deduced from the correlations were considerably altered. In an investigation of the correlations between, and the factors underlying, arithmetic, algebra, and geometry, Oldham[17] has discovered widely varying results in different groups of children. By obtaining introspections she was able to relate these variations to differences in classroom morale, interest in mathematics, the personalities of the teachers, and other non-intellectual conditions.

The imperfect reliability of all intelligence tests has hardly been sufficiently studied. Rodger[18] gave six tests which were closely parallel in content to the same children, and found the average difference between a child's lowest and highest I.Q.'s to be as much as 10 points, practice effects having been eliminated. It would be worthwhile studying these variations by a technique such as Thouless's,[19] to see whether they are due merely to the unreliability of the measuring instrument or rather to "function fluctuation." If the latter was found, it would presumably have to be ascribed to the effects of uncontrolled subjective attitude on test performances.

Unreliablility appears to be greater with younger children, especially those of pre-school age, and this may be explained by the greater difficulty in controlling the subjective situation. The very low level of correlation between tests given during the first few

[15] W. P. Alexander, Intelligence: concrete and abstract, *Brit. J. Psychol. Monog. Suppl.*, 1934, **6**, No. 19.

[16] Discussion at a meeting of the Psychological Section of the British Association for the Advancement of Science, 1937; unpublished.

[17] H. W. Oldham, A psychological study of mathematical ability, to be published in the *Brit. J. Educ. Psychol.*

[18] A. G. Rodger, The application of six group intelligence tests to the same children and the effects of practice, *Brit. J. Educ. Psychol.*, 1936, **6**, 291-305.

[19] R. H. Thouless, Test unreliability and function fluctuation, *Brit. J. Psychol.*, 1936, **25**, 325-343.

years of life[20] might lend some support to the view which Isaacs[21] and Bühler[22] appear to uphold, namely, that the handling of the child's emotional development during these years has a considerable effect on his mature intellectual capacities. There does not seem, however, to be any adequate experimental proof of this. Variations in test performance are found to be greater among uncooperative and emotionally abnormal children, according to Tulchin.[23] No doubt the clearest instances of the effects of emotional states on intelligence are obtained among psychotics, and Spearman admits that neurotic children and adults may sometimes be inhibited from full expression of g.

Now this evidence, while providing good justification for the clinician's unwillingness to separate cognitive abilities from the rest of personality, is not actually incompatible with Spearman's statement. For if the influence of affective states is most marked in the most abnormal cases, and in the youngest children, then it may be much reduced in normal, older children, such as those on whom experiments with objective group tests are usually conducted. We can therefore admit that the psychometrist's results have not been much distorted by his neglect of the subjective situation, but at the same time insist that the application to maladjudged children of tests as objective as those that the psychometrist wishes to use would be futile. There is of course only a difference of degree between "normal" and "maladjusted;" hence many psychologists are likely to continue to believe that truer results with all children would be obtained by testing techniques which are more clinical in nature, although useful approximations are possible from mass experiments with group tests.

Again it should be noted that the greatest effects of orectic factors occur in tests which are not very highly saturated with g (the correlation of Digits Backward, Healy P.C. II, etc. with g tests is likely to average $+0.4$ to $+0.6$); and they are most marked in school subjects and vocational capacities. This confirms Spearman's opinion that "the occupations of ordinary life" are influenced more than "the performances constituting tests." Until better proofs are provided, we would not agree that any tests can be wholly

[20] Cf. N. Bayley, Mental growth during the first three years, *Genet. Psychol. Monog.*, 1933, **14**, 1-92.
[21] S. Isaacs, *Psychological Aspects of Child Development* (London, 1935).
[22] C. Bühler, *From Birth to Maturity* (London, 1935).
[23] S. H. Tulchin, Clinical studies of mental tests, *Amer. J. Psychiat.*, 1934, **14**, 1237-1248.

freed from orectic influences; but the influences may be sufficiently small, among normal subjects, to be dismissed as specific factors or errors which tend mutually to cancel one another, and for the ability aspects of test performances to be isolated and factorized independently.

IV. The Binet Test

Turning now to the Binet Test, we can see that it partakes both of the clinical and the psychometric approaches. It has the flexibility and the control of the subjective situation characteristic of the clinical interview, but at the same time sufficient objective control to yield a realiable measure of ability. Binet's first scale was predominantly clinical; the tests were chosen from children's everyday experience and were intentionally heterogeneous, so as to supply the tester with as many different samples as possible of the testee's mentality. But at the same time the tasks were selected so as to discriminate between older and younger children. To this extent the psychometric conception was adopted; the number of tasks passed provides an index of the testee's mental maturity. In the Stanford revisions, especially in the most recent, Terman appears to have compromised further with psychometric standards, since the tests are somewhat less representative of everyday experiences, their application and scoring have been made more definite, and there has been a more elaborate study of the consistency of each test with the scale as a whole. But this mental maturity has never been regarded as a strictly unitary trait, although it is expressed in quantitative terms, and there is no reason why it should be so regarded. We know, for instance, that the difference between mental ages thirteen and fourteen is quantitatively smaller than the difference between three and four; but these differences are still more unlike one another qualitatively.

The psychometrist naturally objects to the vagueness of this variable and to our inability to define what it is that the Binet, or other individual, tests (e.g., the Merrill-Palmer or performance test scales) actually measure. The M.A. or the I.Q. is a final aggregate score based on tests which refer to a diversity of functions. This summation of heterogeneous items may give a moderately efficient measure of g, because the specific factors tend to cancel one another. But many items probably depend on group factors, or on social background and education, which have nothing to do with g. He would therefore urge either that the test should yield to newer

tests which measure such factors independently, or that, if it is to be retained, the items should be properly analyzed in order to determine the precise extent to which they measure g, verbal, practical or other group factors, educational attainments, etc.

To this the psychologist at a school or a child guidance clinic will reply that his job is the practical one of producing in as short a time as possible as full as possible a report on the intelligence, special abilities, and personality traits of a child; and to predict how these abilities and traits are going to function under various conditions of everyday life. Both in the home, in school and employment, the expression of abilities will be very greatly affected by "insufficiency of will-power, nervous temperament, and susceptibility to complexes." Therefore he does not want the best tests of g or of other factors, but rather tests which most nearly resemble the practical situations of daily life. The fact that g enters more closely into the ability to learn Latin than into any other achievement is a strong point in its disfavor. Any such homogeneous, self-consistent set of tests means a far more restricted set of avenues for viewing the child's mentality. He makes direct use of the specific factors and "errors" which the psychometrist can afford to neglect.

V. Conclusions Regarding the Merits of the Two Approaches

It seems doubtful whether this argument can ever be settled by discussion. The psychometrist can go on to point out that unreliable subjective judgment and intuition vitiate the clinic psychologist's predictions of the child's abilities and traits in daily life, and claim that when factors additional to g are involved in educational and vocational adjustments, they should be isolated and measured objectively. The clinic psychologist can answer that these other tests are, with few exceptions, not yet available; that if they were he would not have time to apply them; and that subjective interpretation and synthesis of their results would still be needed. As it is, he can construct a fairly complete portrait of a child in an hour or two's testing and interviewing, whose usefulness has been demonstrated time and again in the course of his, and his colleagues', experience. If the Binet Test was as inaccurate as its critics suggest, surely there would have been volumes of complaint during the past thirty-odd years from the parents, teachers, doctors, and other clinic workers, who are most affected by its diagnoses. Rather have the complaints been directed at the much more objective group tests.

An experimental comparison of the relative validity of the two approaches will, I hope, one day be carried out. Psychometrists will apply all the quantitative measures of traits and abilities they wish, factorize them, etc. (but exclude completely any subjective interpretation and synthesis of their measures) and make concrete predictions which can be subsequently verified or contradicted. Clinic psychologists will apply to the same children their own individual methods of testing and interviewing and make similar predictions. Judging from a comparison of follow-up studies of cases who received vocational guidance from the National Institute of Industrial Psychology,[24] with the negative conclusions of Thorndike and Lorge's work,[25] I have no misgivings as to which approach would give the best results.

A practical criterion such as this would supply a just proof of the clinic psychologist's methods. But the value of the psychometrist's experiments demands an entirely different type of proof. The criterion of mathematical and logical self-consistency, as at present supplied, is adequate. The present controversy has arisen because the psychometrist has tried to judge the clinic psychologist in terms of his own inappropriate "scheme of thinking," and because the clinic psychologist, looking on the psychometrist's methods and results with his own practical bias, fails to realize its present scientific value and its ultimate applicability to his own work.

In my view, the two approaches should be complementary rather than conflicting; for each type of worker could learn much from the other. Thus the psychometrist has often displayed a narrowness of outlook in the psychological interpretation of his factual results which could have been avoided had he considered them in the light of the more direct and concrete experience of human beings possessed by the clinician. Examples of this are provided by his supposition that precise standardization of the objective conditions of testing will lead to accurate results; by the thoroughly unpsychological theory of specificity which he derives from early experiments on transfer and from the low correlations between character tests; and by his denial that facial, graphological, and other forms of expression can reveal anything about personality merely because his unsuitable correlation techniques yield negative results.[26] He might,

[24] A follow-up of vocationally advised cases, *The Human Factor*, 1937, **11**, 16-26.

[25] E. L. Thorndike et al., *Prediction of Vocational Success* (New York, 1934).

[26] Cf. P. E. Vernon, The matching method applied to investigations of personality, *Psychol. Bull.*, 1936, **33**, 149-177.

on the other hand, render invaluable assistance to the clinical worker, if he would devise more appropriate techniques for testing out some of the latter's "intuitive" generalizations.

The clinic psychologist, for his part, should realize that though practical experience may be a valuable guide, yet it is far from reliable as an authority. The experimentalist can usually provide him with much sounder methods of establishing or refuting the principles upon which many of his methods are based. As a practical worker he seldom has the theoretical interest, or the time, to inquire into the adequacy of his tests,[27] or to follow up and turn to account the discoveries of psychometrists. To be frank, he may either be ignorant of the canons of objective experiment and logical inference, or, even if well trained, may neglect them under the stress of getting immediate results. One especially dangerous, yet prevalent, habit is the attempt to analyze subjectively the component items of the Binet, or of a performance test, scale and to infer from the child's differential performance on these items his differential abilities in various fields. For instance, it is quite unjustifiable to assume, without further proof, that the practical tasks or the memorizing items validly predict any other type of practical or memorizing ability in everyday life. Such attribution of special symptomatic value to particular items or to particular types of response involves the assumption of "faculties" which should certainly be tried out experimentally. This is in fact a case of doing badly what a technique such as factor analysis could do well.

VI. Criticisms of Binet Testing Procedure

With some of the other defects of the Stanford-Binet Test I have dealt elsewhere.[28] The difficulties which occurred when it was applied above the age of about twelve years have been largely overcome in the new revision. But I must discuss two alleged flaws, namely, the lack of objective control of testing conditions, and the excessive dependence on acquired experience.

Now the actual wording of the tests and the concrete material

[27] Cf. note 11.

[28] P. E. Vernon, The revised Stanford-Binet test, *The 1937 Yearbook of Education* (London, 1937). This revision is of course less satisfactory outside the United States, since the translation of the instructions and of the material from American into English inevitably upsets the norms. Nevertheless, it is certainly a better version of the Binet Test than any other available in the British Isles, and I hope that these uncertainties may soon be removed by a British restandardization.

are actually kept as constant as possible; and the clinic psychologist, in his report on a child, does start out from the child's objectively recorded responses. For he realizes that every child should be subjected to the same set of situations if the quantitative and qualitative characteristics which he deduces are to have any significance. But the methods of encouragement, the conversation interspersed between the tests, the order in which the tests are given—these and several other conditions may vary considerably from one child to another. In scoring many of the test items some subjective judgment may enter, though this will be reduced by the explicit instructions supplied with the new revision. Similar variations may occur in the application of performance tests or individual educational tests. And in the summing up of the child's personality and the recommendations put forward, intuitive interpretation plays a large part. Thus in quoting one child's result at the Binet Test, the psychologist may point out that he was nervous and uncooperative, so that the I.Q. is likely to be rather too low, or that, having been previously tested, the I.Q. is rather too high. An allowance may be claimed on the grounds of the child's reading disability, or several of the items in the test may have been arbitrarily omitted because the psychologist thought that the remaining items would give a fairer result. It is not only the inexpert tester who commits these psychometric sins; often the greater the experience, the more the reliance placed on subjective judgment and the greater the number of short cuts or apparent distortions of standard procedure.

Admittedly a great deal more control of such features of testing would be desirable. Yet at the same time most of them are justified by the psychologist's need to penetrate into, and so to control, the child's state of mind. The innumerable minor variations in procedure, and the full indulgence of his own inferences and intuitions, allow him to control the subjective situation *better* than it would be were identical procedures always adopted, as in a group test. Indeed, the reason why a good tester tries to gauge the whole personality from the successive samples presented to him is not so much that he may be able to report on other characteristics besides mental maturity (the dangers of that have already been mentioned), as that his deductions enable him to keep more closely in touch with the workings of the child's mind, and so to insure that he is getting really representative manifestations of the child's intellectual powers.

VII. The Binet Test and Environmental Influences

It is obvious that many of the component items of the test are affected by acquirements such as reading, arithmetic, etc., but there appears to be no good evidence that any other intelligence test of a predominantly verbal type is not also so affected. (Burt's famous regression equation[29] applied only to the 1911 Binet-Simon scale, and might equally well be interpreted as showing the inefficiency of reasoning tests.) If we discount the results obtained from exceptional cases such as Gordon's canal-boat children,[30] we may claim that acquired influences probably play little part in the Stanford Test, since its reliability is known to be very high.[31] Were this not so the hypothesis of the constancy of the I.Q. would not have been so widely accepted. A further piece of indirect evidence may be derived from the Scottish Mental Survey,[32] where it was shown that the average Stanford I.Q. of practically the whole Scottish 11+ population was extremely close to the American norm of 100. Now other experiments such as MacGregor's[33] have proved that Scottish children are very greatly advanced at 11+ over the American standard in educational attainments. Yet apparently this advancement has no effect on their Binet performance. On the other hand, British norms for a well-standardized group test such as Otis Advanced are still distinctly higher at this age than the American figures.

Arguments about nature and nurture in intelligence testing seem rather anachronistic; for surely most psychologists are by now agreed that no test can measure innate intelligence, except in so far as it is expressed through the medium of acquired experiences.[34] It is desirable, of course, to make use of acquirements which are likely to be equally available to the whole population, in order that individual differences in test results may be due to the hypothetical ability behind these acquirements rather than to the efficiency or ineffi-

[29] C. Burt, *Mental and Scholastic Tests* (London, 1921).

[30] H. Gordon, Mental and scholastic tests among retarded children, *Board of Education, Pamphlet* No. 44, 1923.

[31] A survey of experimental results up to 1932 showed the reliability coefficients obtained with individual tests to be just about as high as those obtained with group tests. Cf. C. L. Nemzek, The constancy of the I.Q., *Psychol. Bull.*, 1933, **30**, 143-166.

[32] Scottish Council for Research in Education, *The Intelligence of Scottish Children* (London, 1933).

[33] G. MacGregor, *Achievement Tests in the Primary School* (London, 1934). Dr. W. B. Inglis and I also find Scottish children to be one or two years advanced on several of Burt's and Ballard's English educational tests.

[34] Cf. G. C. Schwesinger, *Heredity and Environment* (New York, 1933).

ciency of training. Yet even information tests have been shown, e.g., by Thorndike's work,[35] to be highly saturated with g. We must admit that a few of the Binet items, e.g., Names of Coins, do not sufficiently conform to this criterion; however, most of these have been eliminated in the new revision.

According to Cattell, the imbrication of the Binet Test with educational acquirements is responsible for the small standard deviation of its I.Q.'s (because education tends to produce a leveling up of differences in ability). But this explanation of the marked difference between the dispersions of the Binet and the group test I.Q.'s is unsound since, as we have seen, there is no evidence that group tests are appreciably less affected by education. Several other possible reasons for this important difference have been discussed by Thomson,[36] but found to be inadequate. The chief reason seems fairly obvious, although it has not apparently been proposed previously, namely, that the more "mechanized" the items of a test, and the more objective the testing conditions, the higher will be the dispersion of the test's I.Q.'s. In the Binet Test, the child is given every opportunity to put forth his best intellectual powers under favorable conditions. Under the much sterner, more formal conditions of group testing, where he has to choose from a set of restricted answers instead of supplying his own responses, the dull child is at a great disadvantage, whereas the bright child is less likely to be affected. Hence the increase in dispersion.

To try out this hypothesis an investigation is in progress at Glasgow, where parallel tests are being applied under individual and under group test (multiple-choice answer) conditions. If confirmation is obtained, then it is clearly unjustifiable to assume, with Cattell, that there exists a "real" standard deviation. We might indeed claim that the Binet Test's figure of 15 or 17,[37] obtained under conditions where there is good control of the subjective as well as the objective situation, is more reliable than a figure of 20 yielded by one group test, or 25 by another, or even 30 by a third. But it would probably be best to agree on some convenient value and to translate the norms of all tests into terms of this value, as Thom-

[35] E. L. Thorndike et al., *The Measurement of Intelligence* (New York, 1927).

[36] G. H. Thomson, The standardization of group tests and the scatter of intelligence quotients: a contribution to the theory of examining, *Brit. J. Educ. Psychol.*, 1932, **2**, 92-112, 125-138.

[37] Terman has now discarded his original figure of 13. With adequate sampling of the population both the old Stanford-Binet (in the Scottish Mental Survey) and the new test yield a figure nearer to 17.

son[38] has already done with his Moray House tests. The majority of psychologists, together with a vast number of doctors, laymen, et al., are already accustomed to a value of about 15 (e.g., they imply this figure when they talk of I.Q. 70 as being the rough borderline for mental defect), and there is no need to attempt the probably impossible task of persuading them to adjust to a much higher value.

[38] *Op. cit.*

MEASUREMENT VERSUS INTUITION IN APPLIED PSYCHOLOGY

RAYMOND B. CATTELL
Psychological Clinic, Leicester, England

I. A Field Where the Controversy Is of Immediate Urgency

Though applied psychology has, on the whole, welcomed the growth of scientific techniques of mental measurement, there has been manifest a resistance on the part of the earlier intuitive techniques to the steady encroachment of objective and statistical procedures. At first sight this resistance may appear to have no more theoretical importance than, and to be based on no very different motives from, that machine-wrecking spirit which has always faced innovations replacing personal skills of which individuals are justly proud. On the other hand, the opposition may be based on a more fundamental disagreement arising from the dissensions of the various schools of psychology. It thus behooves us to examine very carefully the theoretical roots of the opposing points of view rather than complacently to stigmatize these resistances as the rationalization of a natural conservatism.

Since the conflict seems to have come to a head recently in a controversy in the field of child guidance work, I propose to use that incident as a starting point in the present article from which to follow the matter into more fundamental issues of general principle.

The field of child guidance offers an excellent showground for the contending principles because it is flanked by such independent and diverse sources of psychological techniques as psychoanalysis, psychometry, medical psychiatry, and education. At the meeting place of these separate impulses some integration has gone on. For example, the child therapist does not, like the psychoanalyst, consider it possible to work satisfactorily without giving the patient an intelligence test, nor does he, like the teacher, consider it possible to leave out the contribution of the social worker towards manipulating the existing environment. He also looks for relationships between the picture of the child's organic condition, as viewed by the

medical clinic, and the assessment of abilities and interests made by the psychometrist.

Recently a dispute arose at an inter-clinic meeting, representative of the child guidance clinics of England, concerning the best means of measuring intelligence. The present writer took the view that the Binet-Simon and its modifications constitute an obsolete instrument which ignores all that progress in the theory of mental testing which has been made in the last twenty years, and which should be replaced by more up-to-date measures of "g." The opposition was, for the most part, mute or expressed without any reference to the principles of scientific psychology, but the unexpressed resistances of the clinical psychologists soon received academic expression through the able championship of Dr. P. E. Vernon.

II. Criticism of the Binet Test, and Some Rejoinders

Before dealing with Vernon's contentions, I will state the objections to the continued use of the Binet Test. They are:

(1) That its component test items are frequently tests of scholastic attainment[1] and life experience, rather than of "g." The test items have never been validated by statistical treatment except in a few instances,[2] and in these instances the "g" saturation has been regrettably low.

(2) The test items are too few in number to give good reliability and fine grading.

(3) It is inadequate at the higher mental ages.

(4) Certain special factors play a rather large part in contributing to the total score, e.g., the "practical ability" of Alexander's research;[3] "v" and almost certainly the "f" factor of "fluency of association" which is a matter of temperament rather than of cognitive ability.

(5) That if, as most clinical psychologists concede, the test is not concerned with any one ability but a collection of abilities, the attachment of a single quantitative value to this hodgepodge is meaningless.

(6) That in consequence of the contamination of "g" with scholastic attainment, the Binet does not give a scatter of intelligence quotients as wide as that which actually exists. For example, a child of ten with a mental age of six scores more highly than a normal child with an actual age of six. This leads to wrong conceptions of what I. Q.'s are to be regarded as high or feeble-minded intelligence levels.

[1] See Burt's *Mental and Scholastic Tests.*

[2] Hilda Bristol and Raymond B. Cattell, Intelligence tests for mental ages of four to eight years, *Brit. J. Educ. Psychol.*, 1933, **3**, 142-169.

[3] W. P. Alexander, Intelligence concrete and abstract, *Brit. J. Psychol. Mon. Suppl.*, No. 19, 1935.

(7) The personal relationship which arises between tester and testee in this kind of test, and which is sometimes claimed to give greater reliability, is just as likely, in comparison with the impersonal group test situation, to introduce errors arising from shyness and other temperamental variations in the child, as to improve the reliability of the score. Furthermore, the examiner is also likely to be misled by unconcious attractions and hostilities and by the appearance and conduct of the child; for many items are such as cannot be scored with entire objectivity.

Against these arguments Vernon and others have urged that the clinical psychologists are not so greatly concerned to get a measure of pure "g" and that the Binet supplies them with facts regarding the total personality of the patient such as are not given by any test of "g." Similarly, the last report of the Chief Medical Officer of the Board of Education[4] refers to the above criticisms,[5] but "lest school medical officers should feel that they are out of date in using this test" deals with each objection in turn and argues that if the Binet is used in restricted circumstances "little harm can be done." In the end, however, it is admitted that these objections "may vitiate the obtained intelligence quotient as a scientific result, but it makes no differences to the usefulness of the Binet Test as an instrument of diagnosis. There is no other single test in general use by which the examiner can learn so much about the child, his social type, his disposition, his method of tackling problems and his language ability, although the test is not designed to give separate measures of these."

This is a plausible argument hiding a great deal of self-deception. How is one to estimate the separate special abilities if they are not measured as such? What the examiner actually gets by the procedure advocated is a vague composite of ill-defined traits and abilities. If intelligence is worth measuring at all, it is worth measuring accurately, and if we wish to get an estimation of other traits and factorial abilities, they should be measured by ancillary tests, specially devised for the purpose. The way in which these measurements are to be integrated is a separate issue which we shall discuss later. At present, as Vernon himself says at the conclusion of his study of the validity of certain mental tests in a child guidance clinic,[6] "The practising psychologist employs as psychometric and diagnostic tools a number of tests without knowing precisely what

[4] *The Health of the School Child,* 1935. His Majesty's Stationery Office, Great Britain.

[5] Given in Raymond B. Cattell, *A Guide to Mental Testing* (London, 1936).

[6] A study of the norms and validity of certain mental tests at a child guidance clinic, part 2, *Brit. J. Educ. Psychol.,* 1937, **7,** 100.

they measure nor how accurate they are." For example, in most child guidance clinics there is on view an impressive paraphernalia of performance tests, and usually, though the child is given only one intelligence test, he is given three or four different performance tests. What these tests measure is in most cases not known, though where adequate researches have been carried out, it is found that their "g" saturation is very low.[7] The early work of Spearman and his pupils showed that among children having ordinary schooling the "g" saturation of verbal tests is particularly high, but the clinical psychologist, in his concern to escape from what he imagines is an acquired verbal aptitude, runs away from the verbal test into the more devastating error of not testing intelligence at all. The prestige of performance tests among clinical psychologists in the last ten years contributes little to the prestige of the clinical psychologist as a psychologist.

III. Some Consequences of Conservatism in Testing Practice

The prolonged worship of the Binet as a test of intelligence has left us with an encumbering heritage of erroneous conceptions, principally in matters concerning the distribution of intelligence, its development in the individual, and its role in society. In the first place, as all investigators agree, the Binet is not capable of measuring the upper ranges of intelligence distribution among thirteen and fourteen years olds, so that all social, industrial, and clinical researches in which children of this age have been included are inconclusive, valueless, or positively misleading through this distortion in the instrument. Vernon does not defend the test as an implement for research, but only as a means of getting clinical impressions. Unfortunately this distinction has not been recognized by others. If it was, research workers would have to face the alternative of introducing a sounder "g" test into child guidance work or of rejecting from the material of exact research all child guidance records.

Again, owing to the defective scatter in the Binet intelligence quotients, arising from the contamination with life experience discussed above, the majority of psychologists have come to think in terms of a standard deviation of I.Q.'s which is less than the real one.[8] For example, on the Binet the percentage of the population

[7] W. P. Alexander, Intelligence concrete and abstract, *Brit. J. Psychol. Mon. Suppls.*, No. 19, 1935.

[8] Thus, in the *Brit. Med. J.*, Dec. 5, 1936, an editorial notice of the present writer's social survey "Is National Intelligence Declining?" naïvely comments:

below an I.Q. of 70 is not more than 1 or 1½ per cent, but on a newer type of test, verbal, non-verbal or perceptual, administered individually or as a group test, the proportion below 70 is nearer 4 or 5 per cent. There is a similar discrepancy in the upper range. We have to readjust our ideas to an entirely new conception of what is a high and what is a low I.Q.[8a]

Undoubtedly the responsibility for the fact that the intelligence quotient as a measure of mental ability has been going out of fashion in recent years, must be laid at the door of the Binet Test. Where more intensive research should have been devoted to the understanding and perfection of this measuring device—e.g., by investigating the maturation of intelligence with age and inventing mathematical devices to insure constancy of I.Q. over the 14-18 year range; by working out the normal constancy of I.Q. for individuals and different races; and by obtaining an accurate figure for the dispersion obtaining in various populations—many clinical psychologists have instead allowed themselves to be persuaded by the confusion of I.Q. measurements obtained from different tests, into favoring other less fundamentally sound forms of measurement.

Another field in which the Binet Test has probably led to false conclusions is that of genetic psychology where Freeman and others, supposing the test to measure intelligence or "g", have arrived at results giving a quite inadequate allowance for the part played by heredity.

The expenditure of time and money over twenty years on improving, standardizing, and applying an instrument which was never rooted in adequate research has deflected a stream of valuable effort which should have gone to newer problems. This charge is made initially with regard to the influence of the Binet Test upon the progress of intelligence measurement, but it holds even more strongly with regard to the testing effects upon the growth of techniques for

---

"The results of applying special (non-verbal) tests show a greater variation of intelligence quotients than any previously reported (i.e., on Binet Tests). This suggests that the tests are either unreliable or not satisfactorily standardised." The standard deviation given by this non-verbal group test is 21.9. The standard deviation for the Binet originally worked out by Terman was 13. As Vernon points out, most English psychologists soon came to accept 15 as the standard deviation of the revised Binet, while that for the latest revision is nearer 17. Surely it is not too much to expect that with further and more complete revision the Binet will approach the 20-25 figure given by newer tests carefully selected to test "g" in as pure a form as possible?

[8a] The recent Scottish Mental Survey found a scatter of intelligence quotients somewhat larger than that to which we have become accustomed, but as, unfortunately, the test used was first standardized against the Binet, even this new scatter gives no indication of the real standard deviation.

assessing temperament and character. The examiner imagines that he can deduce a good deal about the child's disposition and temperament from reactions in the Binet Test situation and in performance tests. There is, however, no proof that the qualities of persistence or impulsiveness shown in these miniature test situations—which themselves do not necessarily have any relation to the child's total conative life pattern—will throw any light on his persistence or impulsiveness in the major situations arising in school and home.

And if the examiner is judging by conversational reactions or by facial expression, these can be observed equally well with any other type of test or with no test at all. Indeed, it becomes increasingly clear that the "subjective judgment" which the clinicians hold in such favor, and which must be used to modify test results, has no connection whatever with the test procedure itself but acts, so to speak, as an independent vision through the interstices of the test.

But these methods give even less reliable measures of noncognitive traits than the Binet does of "g." As the studies of intuitive judgments and interview assessments show, such methods are far less valid than the interviewers have always so confidently imagined and certainly less valid than the as yet quite rudimentary temperament tests which have been devised by psychologists.

A test of "fluency of association," as a measure of surgency-desurgency of temperament, offers a validity (uncorrected for functional fluctation) of about .6, but no estimate based on an interview of the same length of time, with or without the use of the Binet or a performance test situation, will give the same validity. And the perseveration test, for all its low consistency, gives indications of character and temperament structure which are well-nigh indispensable in clinical diagnosis and treatment. The finding that the high perseverator will be relatively difficult to treat is borne out by clinical experience. Again the prediction, from the test alone, of the type of instability to which the low perseverator will be subject, has time and again proved to be uncannily accurate where the intuitive judgments of those who met the child in the clinic have indicated only a stable and attractive type of personality. That these tests are not more highly developed in consistency and validity arises from their neglect by the clinical psychologist under the complacent but entirely false assumption that the conviction with which he makes his intuitive judgments has any relation to their real reliability.

IV. Defences of the Intuitionist Standpoint

Let us now turn to certain weighty objections which have been urged against the psychometric position by the intuitionist, for although some of them will prove to be sheer misunderstandings, others cut right down into the philosophical basis of mental measurement.

(1) Against the factorial methods of analysis of personality, Vernon has urged that the "statistical approach is cross-sectional; the clinical and vocational is longitudinal." But would it not be true to say that without accurate cross-sections, no science of longitudinal changes could arise? One could use a cross-section alone in a variety of ways, but one cannot speak of the longitudinal study of personality at all, i.e., of the change of personality with time, except on the basis of cross-sectional studies (though longitudinal studies may afterwards be used to check up on cross-sectional methods).

(2) The practical objection has been raised to the factorial testing of special aptitudes, disposition, and temperament as such that the time allowed in clinical work is insufficient to cover the ground. It has been suggested that it would require ten hours to make a complete inventory of the individual's special aptitudes and temperament, whereas a clinical psychometrist has, at the most, two hours to give to each individual. Furthermore, except in a few instances, the tests for these special and general factors, other than intelligence, are not available. But the improvement of techniques, now that the main difficulties are overcome, is only a matter of more time given to research; the solution of one problem is the solution of the other; for the more perfected the test, the shorter will be the time required for an accurate measurement.

(3) It is objected by those who think that the personality can only be assessed as a whole that the measurement of any trait or factor in isolation is of no practical use or importance. Now although the measurement of a trait will not tell us anything about the part it plays in the total personality, it suffices, under many special conditions and in proportion to its accuracy, to permit a prediction of the person's behavior in some limited field. For example, the performance in a group test of intelligence leads to reliable predictions of success in other tasks requiring mainly "g" under similar working conditions. This fact has its greatest practical importance in the techniques of vocational selection, where, for

instance, it may be desirable to select for one occupation by manual dexterity test alone.

A refinement of this objection is the admission that it may be desirable to measure a trait but not a factor. Thus Thomson argues against "the disadvantage of introducing unnecessary middlemen between the test and the occupational criterion."[9] Vernon expresses the same objection[10] in a wider context, saying that the "psychologist does not want the best tests of 'g' or of other group or specific factors, but rather tests which more closely resemble the practical situations of daily life." On this view the Binet Test, for example, is a better test because it measures, in addition to "g," other factors which influence the effectiveness of "g" in certain situations.

The error of this view would seem to be in the assumption that "the practical situations of everyday life" are all of a kind. But in reality the particular combination of factors which leads to success in the Binet may be quite inapt to many of the diverse situations in which the individual has to use his "g." By having the cognitive and conative elements of personality at his disposal, instead of crude combinations of unknown composition, the psychologist can re-combine them to give a truer picture of the personality as it functions in each particular performance or situation, in this way showing a far truer appreciation of the lability of the living personality than is shown by those who begin by denying the psychometrist any appreciation of so subtle a quality.

Admittedly when the psychologist has very limited practical objectives it may be convenient sometimes to work with a composite score, as the chemist, in commercial syntheses, may not find it necessary to purify his raw material or break it down into simpler elements beyond a certain point. Experience alone can decide which is the more reliable procedure in any given situation, and our practice is bound to vary as the purposes in view alter. For example, in the vocational guidance scheme of the Leicester schools, it is desirable for the records to contain as full an analysis as possible, in terms of factors, of each personality; for the individual may apply for a variety of jobs over the course of years, each of which requires a different combination of qualities; but where the fitness of a series of persons for one particular job is alone in question the direct occupational criterion has the attraction of simplicity. One cannot

[9] Godfrey H. Thomson, The factorial analysis of human abilities, *Human Factor*, 1935, **9**, 180-186.

[10] See pp. 106 f. of this journal.

help feeling, however, that the passion for *ad hoc* vocational tests, with all their meretricious simplicity, which has characterized American as opposed to English vocational psychology, is largely responsible for the debacle revealed by Thorndike's investigation of the reliability of vocational advice schemes. But even when the unanalized test proves to be practically valuable, its blindness is aesthetically unsatisfying to the scientific mind.

(4) It is claimed that the test which allows some play to subjective judgment in deciding the assessment of, for example, intelligence, is in fact more accurate than one with rigidly standardized objective procedure. Greater reliability is said to arise, in the first place, because the test is more pliant—"the interspersed conversation, the methods of encouragement, the order in which the tests are given . . . . may vary considerably from one child to another"—and, secondly, because the examiner can make allowances when the child has failed for some irrelevant reason.

No experienced clinical psychologist would object to the first of these procedures, providing that it applies only to the right preparation of orectic conditions in a cognitive test or of cognitive conditions in an orectic test; and a well-designed test should certainly permit adaptation to circumstances (e.g., a group test should permit of short sittings or long ones), for it is a mistake to think that "scientific" or "objective" are synonymous with "rigid" or "foolproof."

But the nature of the freedom which the examiner can permit himself, particularly with regard to the second procedure, must be determined by scientific considerations and knowledge of factor psychology. For example, Vernon criticizes the newer tests of "g" because the subject "has to choose from a set of restricted answers in limited time instead of supplying his own responses at his own speed," but if speed in higher cognitive performances is of the essence of "g" the longer time allowance arbitrarily allowed the duller child in responding to the Binet Test is bound to give a falsely inflated I.Q. Similarly the fluency favoring a high score in inventive answers is a temperamental quality having no relation to "g." In short, the variability of procedure and interpretation must have a definite relation to the qualities being measured and must be based on a scheme of scientific knowledge not on immediate intuition.

(5) Principally from the medical world comes criticism of the following kind: "The statistical psychologist is interested primarily

in the trait rather than in the patient, in mathematical validity rather than in clinical usefulness, an attitude which is justifiable in the laboratory but not in the consulting room,"[11] and again "If the intelligence and capacity of an individual can be calculated in terms of an hard and fast mathematical formula, then we have no longer any elasticity left in assessing these. This is where we fear the results of all these elaborate tests are apt to prove fallacious."[12]

The implication seems to be that there is some other and esoteric standard by which the clinician can judge the validity of his tests. The clinician, we are told, depends on "personal experience, or the experience of colleagues rather than on scientifically established facts." What is this other standard but the complacent self-delusion which comes with intuitive judgments frequently made and checked by no more reliable objective record than that which the visitor to a fortune teller applies to his memory of the predictions made? A little study of the researches on intuitive judgments should do something to bring humility to the authors of such dogmatic statements.

(6) It is urged that the psychometric approach can never handle the total assessment of personality since it fails to take account of qualitative differences of mentality. This, of course, is untrue. The psychometric method deals with quantitative differences only in traits of the same quality. It distinguishes quality before it concerns itself at all with quantity. It does not measure a verbal quality of intelligence against a mechanical quality but takes care first to isolate the unitary qualities in which measurement is to proceed. Perhaps some such term as "dimension" would cause less misunderstanding than "quality," but both are to be regarded only as metaphors for a manifoldness in personality which may some day receive more technical definition.

(7) Vernon reminds us of the disconcerting fact that different research workers have obtained different special factors and that, according to Thomson, this must always be so because an infinite number of different factorizations of the same test results are possible.

The issue touched upon is too great to be dealt with adequately here, though it is obviously of fundamental importance to the remaining arguments in the article. While the analysis of any table of correlation coefficients is at present in part arbitrary, the analysis of a hierarchical table into general and special factors

[11] Review of "A guide to mental testing," *Brit. Med. J.* 1937.
[12] Review of "A guide to mental testing," *Med. World*, 1936.

remains on the same foundation on which it has always stood—namely, that the analysis satisfies the principle of parsimony of hypotheses better than does a sampling theory.

Until the large scale research by Thorndike, Holzinger, and others on Unitary Traits has come to fruition it is premature to say that disagreement as to factors found in small-scale researches is due to anything but sampling errors, and even the discovery that factors are relative to conative patterns and particular social groups would not destroy their usefulness for individual psychology within those groups.

(8) It is alleged that concentration on the objectivity and scientific validity of tests of ability as such causes the clinical psychologist to neglect the more important conative and emotional elements of personality, without which the abilities are meaningless.

There are two distinct issues raised here, and the first is quite a simple one. It is exemplified by the statement that the measurement of intelligence is itself inaccurate in some of the newer tests and in group test situations because the examiner cannot be as certain as he can in the Binet Test situations that the candidate's full attention and effort is being given to the exercise.

Now, in my opening criticisms of the Binet Test, this defence has been anticipated by pointing out what is not so commonly realized: that the personal interview situation of the Binet Test may frequently be inferior to the impersonal group test situation by reasons of shyness, emotional disturbance or contra-suggestibility in the subject, quite apart from emotional changes in the examiner. The conversational situation brings in stronger, but also more confused, conation to a type of performance in which strong conation is in fact unnecessary for a correct estimate of cognition. All systematic researches have shown again and again that the strengthening of effort beyond a certain relatively low level does not improve "g" performance and may positively impair it.[13]

Some discussions on group test validity would lead one to imagine that such tests are only given to classes of children in a state of uproar: the verdict of experiment is that, at least for older children and adults, the mental set in a group test situation is one which gives the optimum intensity of effort for the measurement of "g."[14]

[13] See Spearman, *Abilities of Man* (London, 1937), p. 333. Vernon expresses the upshot of researches by saying "the greatest effect of orectic factors occurs in tests which are not very highly saturated with 'g,' i.e., in the 'engines' not in the 'mental energy' itself."

[14] Because the controversy has been particularly associated with the Binet Test,

The second issue raised by this objection that the temperament-character factors are neglected is of a more subtle kind. It is the suggestion that the cognitive special factors cannot be separated from emotional drives, and that the treating of them in isolation as cognitive special factors ignores the convincing clinical observation that the special abilities in which they occur are frequently affectively determined, belonging to the emotional rather than the cognitive realms of personality.

Now although the S's—specific and group factors—which operate with "g" are commonly considered as cognitive in nature, the term "general factor" being applied to such new factors as "f" or "p" covering a different dimension or modality of personality, they need not be so in origin. No one has proved or even asserted that such S's as operate in mechanical aptitude or verbal ability are constitutional and inborn abilities in the sense that "g" is. Many S's are undoubtedly subject to environmental influence and in clinical work one can often perceive how a special ability has developed to an unusual degree as an overcompensation or reaction formation in the adjustment of a particular temperament to the given environment. Special abilities grow or atrophy as part of the individual's total organic adjustment to the situation. Observation suggests, for example, that a marked mechanical aptitude may be acquired by a sensitive but assertive child who, over the course of years, has turned from social expressions of libido and has no great verbal opportunities or interests.

If "cognitive" abilities and memories respond to emotional determination over long periods in this way it is at least just conceivable that they also function or fail to function in a more momentary fashion in response to changing conscious and unconcious motivation. The analysis of abilities into "g" and S's makes no assumption that these S's operate in academic detachment from the emotional roots of personality: it investigates the connections statistically

---

some participants may fall into the error of supposing that the psychometric versus intuitionist issue is bound up with the group versus individual test discussion, but of course this is not so; it is only that the individual test situation gives scope for the vagaries of intuition, whereas the group test situation does not. A good deal of further experiment is urgently needed to decide whether the objective test is more accurate in a group or individual test situation, and in any case it is time that the claims of the individual intuitive test to greater reliability were put to the verdict of experiment. A recent experiment in which 5/- was offered to every boy who could improve on his previous score failed to yield any significant difference in the group average I. Q. (H. H. Ferguson, Manchester Meeting of the Brit. Psychol. Soc., 1937).

with a view to revealing in the end how the emotional and the cognitive elements do in fact interact.

(9) The most interesting criticism, and one that points to a very real problem, asserts that when the various special factors and dimensions of personality have been sorted out and accurately measured the examiner still knows less about the personality than he would through knowing it intuitively as a whole. Vernon quotes Myers that "man is more than a mosaic of separate abilities" and Terman, who says, "The total is an organismic, not an additive total." But who ever suggested that the elements should be combined additively? Surely no one but these same clinical psychologists, who, in fact or by implication, add together mental ages on diverse performance tests to get a picture of the individual's "mental level."

At this point we have to reckon with all that goes under the name of gestalt psychology. Since the general relation of both factor psychology and noegenesis to the doctrines of gestalt has already been treated by Professor Spearman elsewhere,[15] and most illuminatingly with regard to the shortcomings of the gestalt principles in the matter now under discussion,[16] I will confine myself to a few suggestions only.

One feels that in this matter the gestalt psychologist contributes nothing except a magic word, the purpose of which is to cry "Halt!" to the irreverence of further investigation of personality. If he belongs to that school of gestalt which admits the priority of elements he yet makes a great business of the "integration" of those elements. What a wealth of ignorance is hidden by that word "integration"! Instead of leaving everything under the blanket of a portentous word, the factor psychologist desires fully to investigate the separate factors in order to reach the second stage—that of discovering their manner of interaction.

For even a gestalt, in the form of a curve or surface, can in general be represented as the function of independent mathematical entities, and it is conceivable that the total personality can likewise be defined in terms of the values for the separate dimensions or factors. The riddle of creative synthesis is presumably a riddle only so long as we are unaware of the full properties of the synthesizing elements; and the saltness of common salt would be deducible to

[15] The battle between "intuitionists" & "psychometrists," *Brit. J. Psychol.*, 1934, 24, 403-408.
[16] *Ibid.*

a physical chemist understanding the full nature of sodium and chlorine.

The truth of this view is not certain, but to suspend research on the ground that the opposite view may be true is to convert science into a paralyzed and pessimistic philosophy. Surely the task of psychology is to proceed as if a more hopeful view were true until that view is proved empirically to be untenable. There are any number of partial relationships between elements in the total of personality which can for certain be profitably investigated immediately. In what way, for instance, are the values for "g" and the group factor of manual dexterity and a specific dexterity factor to be combined? What is the influence of high or low "p" upon the expression of "g" through some group factor in memory?

In seeking to answer the questions that arise it is possible that we shall have to think of the measurement of psychological powers not in biological terms—i.e., through comparisons with the reactions of other individuals of the species, as with the I.Q.—but in "interactionist" or "cosmic"[17] terms of real effects of these powers upon the social and physical environment. A momentary digression may make these terms more clear. We realize that the measurement of the power of a muscle or even of the whole body might be expressed by percentile or quotient units in relation to other similar muscles of human bodies; but that alternatively it might be expressed directly in terms of the physical work which can be done. In the case of mental functioning the terms in which the objective transactions with the external world can be expressed are not so immediately available in units ready made by another science, but doubtless they exist.[18] It is in connection with the measurement of interest that this "cosmic" means of assessment first suggests itself, and I have dealt with that question fully elsewhere;[19] but both this method of measurement and its application to expressing the integration of factors can perhaps be best illustrated here by the research[20] in which it was found that in a given community the number of people known to any individual (and the number who knew him) was (roughly) proportional to his score in an "f" test. When next the social

[17] I use "cosmic" for lack of a better word and because it has already been used, as opposed to "solipsistic," in the field of interest measurement, where it is more apt. The term "interactionist" has too many other associations.

[18] The field of economics suggests itself as one place in which they might be found.

[19] The measurement of interest, *Char. & Person.*, 1935, **4**, 147-169.

[20] Friends and enemies: a psychological study of character and temperament, *Char. & Person.*, 1934, **3**, 52-64.

consequences of the "p" dimension of personality were examined it was found that these numbers were not determined without taking into account the "f" score at the same time. High "p" with high "f" produces a personality creating far more than the average number of enemies; low "p" with high "f" resulted in a personality of great popularity. With low "f" the effects of "p" were barely noticeable.

This example illustrates that the consequences of the integration of the elements of personality can at any rate be expressed in terms of "cosmic" measurement; in this case by the actual number of people attracted or made hostile. Such a measurement expresses the actual transaction that the individual mind has with its environment. It is too fanciful a hope that such measurement will eventually be expressible in terms of energy transactions and indeed of mass-space-time, thus opening the door to a principle of conservation of energy which will include mental energy (though not necessarily on the plane of simple convertibility).

V. Wider Implications of the Controversy

To treat this controversy as if it were merely a difference of opinion between scientists—a disagreement of the standard scientific kind leading to solution by scientific method and reasoning—would be to blind oneself to the fact that issues of a different and indeed of a more fundamental kind are actually at stake. Until these are consciously separated out, the purely scientific solution will continue to be confused and delayed.

In the first place the purely scientific issue becomes sidetracked because the clinical psychologist wishes to treat his techniques as an art rather than as a science. Admittedly the use of scientific instruments must always remain, in the highest flights of diagnosis and treatment, an art; but only in arts at a very low stage of development do we find the instruments themselves based merely on intuition and rule of thumb. The painter's colors become products of science, the physician's drugs become chemicals of known physiological effect; it is only in the combining of these and their nice adjustment to the demands of each particular situation that skills of an artistic kind have their place.

Without a sure scientific basis for the tests used in child guidance work; without a clear understanding, through statistical researches, of the significance of separate measurements, the ways in which test results may be combined, and the manner of their

variation with time and environmental influences, no art worthy of the name can be developed. Of what value is it for individual clinical psychologists laboriously to build up artistic skills and intuitive lore around crude instruments that are being momentarily superseded? To make an incommunicable art where one could make a universal science is, at the least, shortsightedness.

Many who would agree with the above as far as diagnosis and the understanding of personality are concerned, believing that fuller and more reliable knowledge is to be attained through objective test procedures than through direct contact of personalities, will yet argue that where treatment is concerned detached understanding is not enough; the psychotherapist must know the patient intuitively because he has to treat him partly through the pulls and thrusts of the patient's personality upon his own. But this is true only of certain kinds of psychotherapy and particularly of those in which suggestion and transfer play a large part. In much child guidance work the psychotherapist himself need not come into the picture. Instead, from his understanding of the case, he can bring about that rearrangement of the more powerful and permanent forces of the home and school environment which he knows will favor the desired readjustment of personality. Even where the psychologist himself is the chief pivotal force brought to bear, his understanding of the case in the objective terms of some psychometric or psychoanalytic formula does not prevent his translating that understanding into intuitive terms, while it does insure his making that allowance for his own personality make-up which would not necessarily be made by the intuitionist. In short, psychology as a practical art is no less of an art, but rather a far more refined art, through being based on a more extensive scientific foundation.

This divergence of practices having been dealt with, we can turn to the ultimate cleavage in theory, which is one running far beyond the boundaries of natural science. For, as Spearman has pointed out, the tenets of much Struktur and gestalt psychology are nothing less than a denial of the validity of scientific method. Still more remote from scientific methods and concepts is the intuitive and descriptive characterology of Dilthey, Klages, and Spranger, but these writers do at least openly acknowledge their apostacy by flying the flag of the *Geisteswissenschaften.*

When Klages, for example, from purely personal insight asserts of character qualities that "The attribute is never a part of the

concrete thing, it is only an aspect thereof," and "All kinds of properties, whether of objects or of human characters, can indeed be distinguished but not separated," we are not surprised that he is forced to the position that "Enquiries into the human mind, because they are entirely concerned with the realm of qualities, are in their methods and aims fundamentally distinct from what have so far been called the natural sciences. The aim of the latter is to reduce the world, in abstraction from its appearances, and for the purpose of making calculable the associated special processes, to a system, in which there are finally only two kinds of relationship—movement and quantity."

Beyond this point the philosophical leaning which has always characterized German psychology develops into an unmistakable mysticism. And it is from this source that sallies, like those of nomads into agricultural regions, are constantly being made upon the cultivated field of scientific psychology. One cannot avoid thinking that those within this field who rise in sympathy are impelled by a desire to escape from the irksome restrictions of scientific reasoning. The intuitionist movement is not an organized attempt to make scientific progress on new lines but a disguised desire to escape from the rigors of science altogether. It is surrealism breaking out in the most susceptible of the sciences.

But whatever its origin, a consideration of epistemology will show that it is foredoomed never to reach its goal. The intuitionist desires to know another personality directly, by empathy instead of in an objective fashion through the mediation of a mathematical formula of personality. Because the mind of another human being is more akin to his own than is, say, the mind of an insect, the life of a plant or the metallic quality of brass he is misled into believing that he can apprehend it directly in a way which he would never dream of claiming in respect to these other qualities. Yet in fact he is equally tied by solipsism, or, if he wishes to proceed as far as other philosophies will permit, he is compelled to know the other mind as a phenomenon not a noumenon. He is in no different position from the physicist attempting to understand the nature of electricity.

It is the especial danger of psychology that the psychologist is encouraged to use intuition because by reason of the partial similarity of minds, it yields an attractive proportion of correct guesses in predicting the thoughts, feelings, and behavior of others. But

even in practice its use leads in the end to errors more impressive than its successes and to viewpoints which, though similar, permit no machinery of reconciliation. Thus Jung maintains that Freud's theories of personality are merely a reflection of Freud's own personality structure falsely projected on to many minds that are in part different. And every intuitive judgment made in a psychological clinic is a similar stereotype distortion of a unique mind. Intuition has an indispensable place in research, as a scaffolding under the shadow of which objective investigations may be built up; but propounded as an independent method of arriving at psychological knowledge, it would seem to be a pure illusion.

CLOUD PICTURES: A NEW METHOD FOR TESTING IMAGINATION

WILLIAM STERN
With the Collaboration of Jean MacDonald
Duke University

The question of methodology in the field of imagination meets with particular difficulties. Imagination, being a spontaneous activity relatively free from objective control, cannot be studied under the sharply determined test procedures and in the unambiguous task situations which are commonly employed in tests of intelligence, memory, attention, and the like. Since freedom is the kernel of fantasy, the chief methods in the study of imaginative activity consist in the observation and interpretation of free expressions of imagination in play, art, dreams, and everyday life. Nevertheless, there is also a need for tests of imagination that may serve as supplementary tools of theoretical research and as indices of personality traits. The elaboration of such tests requires that the task offered from the outside shall have the character of a fantasy stimulus allowing freedom for the realization of internal tendencies. Although the test performances are basically reactions, spontaneity within reactivity must be taken into account. The contribution which such tests may make must proceed from the evaluation of qualitative features and typological differences rather than the measurement of quantitative grades of performance. Despite the absence of quantitative exactitude, it may be expected that important insights into the manifoldness and complexity of mental data may be acquired by such means.

There are two basic kinds of fantasy-stimuli that can be offered in tests of imagination, the linguistic and the visual. Visual stimuli may be divided further into "meaningful" and "meaningless" forms. In offering meaningful pictures, we require the subject to interpret and supplement imaginatively the story presented. With meaningless forms (such as may appear in cloud conformations, ink-blots, geographical outlines, and other accidental shapes) we

examine the subject's ability to discover meaningful objects and scenes within the stimulus matrix. It is with this last method that our considerations deal exclusively.

The best known material consisting of meaningless forms is the Rorschach Series.[1] This series is made up of ten cards upon which blurred symmetrical forms have been created by folding a white paper upon which black or colored ink had been spilt. The Rorschach Series enjoys a very extended application to problems of psychopathology in many countries. Indeed, in certain quarters, the method has tended to acquire a monosymptomatic despotism, which must always challenge the criticism of unbiased psychologists.[2] But appropriate as the material may be for the purposes of psychopathology, for which it was intended, its adequacy to general psychological problems, especially those of imagination, may be criticized on several counts. Two of the most important defects of the material are these: (1) the outlines of the patterns are sharp and unmistakable; (2) the configurations are wholly symmetrical. This narrowly restricts freedom of imagination. The perception of symmetrical forms, insects, butterflies and bats, anatomical crosscuts, and pairs of people standing face-to-face, tends to be overemphasized.

Cloud-Picture Method

In the Hamburg Institute an attempt was made to design fantasy-stimuli which would avoid the disadvantages of the Rorschach Series. Following my advice, Karl Struve elaborated a series of three pictures which resemble clouds rather than ink-blots.[3] Picture A offers the most distinct shades and forms. Picture C is the vaguest, and Picture B occupies an intermediate position (See Figures I-III). There is no symmetry or other regularity in these pictures. Moreover, instead of sharp outlines, salient upon a white background, the Struve pictures show irregular spots of different shades of black and gray, more or less blurring into one another. Thus it is left to the subject to single out the items to be interpreted, i.e., to determine what parts of the picture seem to

[1] H. Rorschach, *Psychodiagnostik* (Bern and Berlin, 1932).

[2] For recent treatments of the Rorschach-Series see: S. J. Beck, Introduction to the Rorschach method, *Res. Monog. Amer. Orthopsychiat.*, 1937, No. 1; S. J. Beck, Problems of further research in the Rorschach test, *Amer. J. Orthopsychiat.*, 1935, **5**, 100-115; P. E. Vernon, The Rorschach ink-blot test, *Brit. J. Med. Psychol.*, 1933, **13**, 89-118; 179-200; 271-291; also M. R. Hertz, The Rorschach ink-blot test; historical summary, *Psychol. Bull.*, 1935, **32**, 33-66.

[3] K. Struve, Typische Ablaufsformen des Deutens bei 14-15 jährigen Schulkindern, *Zsch. f. angew. Psychol.*, 1932, **37**, 204-274.

represent the outlines of single figures. Under these conditions, imaginative activity has greater freedom to express its own tendencies.

The freedom of activity thus permitted the subject demands a new experimental procedure. Whereas Rorschach requires only oral reports of what the subjects *see*, the Struve technique requires a registration of the outlines projected into the picture. Each subject receives his own copy of the picture upon which he is asked to trace the outlines of the alleged objects and to label them in the margin. This procedure offers a permanent record of the child's imagination which supplements and clarifies the protocol of his words. In addition to the drawing-record of the fantasy-images, a careful protocol is taken not only of the words but also of the total behavior of the subject, his handling of the picture, his expressive movements, and the like.

Previous Experiments

Let me describe first some experiments carried out by Struve in group and individual testing of fourteen- and fifteen-year-old children. Our own experiments with an improved method and enlarged aims will be discussed later. Struve, like Rorschach, largely neglected the problem of the content of imagination. His purpose was to determine the characteristic *forms* of perceiving and interpreting and to investigate whether these forms of the imaginative process correspond with more general features of intellectual activity.

The data acquired from group experiments with 223 subjects enabled Struve to list five typical *modes of interpretation.*

a) The "enumerative type" *(aufzählender Typus).* The children pick out single disconnected objects referring to a narrowly limited realm of experience. No attempt is made to synthesize the items.

b) The "explanatory type" *(erklärender Typus).* Single objects tend to be conceived abstractly and their relations interpreted logically. Often the subject attempts to display items of academic knowledge.

c) The "intuitive type" *(anschaulicher Typus).* Total views are developed in a plastic and realistic form. Sometimes the phantasmas resemble eidetic images.

d) The "inventive-unregulated type" *(erfindend ungeordneter Typus).* This type is characterized by uncontrolled invention, unrealistic connections and goalless fancy.

e) The "inventive-regulated type" *(erfindend geordneter Typus).* Total situations and stories of original invention and superior goal-directedness are created.

One case belonging to the "inventive-regulated type" may be selected from Struve's article. The subject was a fifteen and a half year old boy who was tested individually. The protocol shows, as the drawings and marginal notations alone would not, the temporal development of the interpretation. An interrelated process of thinking and imagining is revealed. A visual impression first evokes a concrete image that is immediately related in thought to a meaningful context. From this meaning nexus, the later visual forms acquire their specification. Occasionally a new image appears which fails to fit into the theme of the fantasy and necessitates a reinterpretation of the previous forms. The total meaning continually increases in structure and definiteness.

Seldom have psychologists had the opportunity to observe the lively processes of imagination so clearly. The protocol is quoted, therefore, almost in full. Explanatory remarks for the reader are given in parentheses.

Subj. "I see a man. Here's a boat and a dog."

Exper. "Won't you write them on the margin?"

Subj. "Not yet. There is more. Perhaps they belong together.—The boat is a canoe. Here, above, are faces in the wood; they are ghosts. These people are Indians. Here is a man with a cap; it is a white man. Perhaps he is a teacher of the tribe. No, he is not. (Apparently because it does not fit into the dawning context.) Here is somebody lying down. He is already dead. There is a struggle between Indians and white men. The dog is watching the canoe. The ghosts in the background are the idols of the tribe. (Note this change of interpretation.) They are to help the struggling Indians. Here are blocks of stone, and water in between them. They have come here in the canoe. This is another idol,—with such a big nose, too, and a large mouth. Here is a man; he looks like an Indian. He is going to attack the white man from behind. Perhaps the whole scene occurs in a wood,—or rather, in a subterranean grotto."

Exper. "Why?"

Subj. "Because here are rocks and water. And they have hidden their idols here. (The new conception of the grotto leads immediately to an interpretation that explains the idols and the struggle within the grotto.) Here is another man with a hat. (This figure was interpreted in the beginning as a teacher. Now he is inserted into the context.) I suppose he is shooting from an ambush. But perhaps he is helping the hurt people. . . . (Finally, as if he would sum up the results of the whole process, he ends with these words:) Title, 'Struggle for idols in a grotto'."[4]

Recently Jacobsen[5] published another investigation in which Cloud pictures were employed. He introduces the important cate-

[4] *Ibid.*, pp. 231-232.

[5] W. Jacobsen, Charaktertypische Arten des Deutens von Helldunkelbildern, *Zsch. f. Psych.*, 1936, **140**, 86-108.

gory of *Formgebundenheit,* i.e., the degree to which the outlines drawn by the subject comform to the forms objectively presented in the picture. Jacobsen tested twenty boys of about the age of eleven years. He divided this group into Schizothymes and Cyclothymes according to Kretschmer's typology, and found, with the Cloud-picture test, certain characteristic differences in imaginative attitudes. The Cyclothymes occupied, almost exclusively, a broad middle range of *Formgebundenheit,* while the Schizothymes were to be found at both poles. Evidently the "Schizothymic type" is not unique. There are Schizothymes who experience the pressure and power of the external world so overwhelmingly that they do not allow themselves the faintest deviation and stick rigidly to the forms given. There are others who neglect and despise the claims of reality and indulge in autistic confabulations. The objects seen and drawn have little consistency with the external stimulus.

Present Form of the Method

In the Section for Child Study recently added to the Department of Psychology at Duke University, the Cloud-picture method has been improved and applied to approximately fifty nine-year-old and thirteen-year-old children. In this work I have enjoyed the valuable collaboration of my assistants, Dr. Wally Reichenberg (1935-36) and Miss Jean MacDonald (1936-37), and the gracious co-operation of Mr. W. F. Warren, Superintendent of the Durham Public Schools.

The aim of the new investigation was twofold: (1) to give the method a standard form which would be applicable to general and comparative problems as well as to individual psychodiagnosis; (2) to compare the imaginative activities of white and colored children.

The three Cloud-picture designs were retained in the form elaborated by Struve. The technique of their presentation, however, was made uniform through a temporal segregation of the different activities required of the subjects. A test of suggestibility was added to examine whether or not imaginative processes may be influenced by suggestive stimuli. The following is an exact description of the method as formulated by Miss MacDonald:

Experimental Procedure

The experimental hour is divided into three periods:

(1) The first period begins with a statement of the task: "Look at this and tell me how you like it." If the subject makes no response, the

experimenter asks the more definite question: "What do you see in the picture?" The subject is then left to *report* spontaneously the various items and scenes which he finds. The experimenter meanwhile protocols the report exactly. If the subject does not offer to turn the picture in different positions, he is told that he may do so. No constraint is introduced.

(2) The second period is given over to the *drawing-task*. The subject is asked to outline the various items and scenes which he has mentioned during the first period, and to write the names of the items in the margins of the picture. No prompting is given unless the subject fails to include all of the items previously mentioned. The number of items on which prompting is required is recorded.

(3) The third period is devoted to the test of *suggestibility*. Three suggestions are made, each in the following form: "Some people have found a —— in the picture. Do you see it?" Three classes of objects have been used for these suggestions: (a) asymmetrical familiar objects, as dog, cat, etc.; (b) symmetrical familiar objects, as bat, butterfly, etc.; (c) angular familiar objects, as church, table, etc. Since the forms in the pictures are asymmetrical and non-angular, the three classes of objects, in the order given, are assumed to impose increasing strain upon the subject's suggestibility.

A time report is added in the protocol. This includes the time for each of the three periods, the total time for the experiment, and the time, in seconds, for the acceptance or rejection of each suggestion.

The three pictures are presented separately at intervals of one week. The procedure is the same on each occasion.

Schema for Evaluation

A tentative schema for the evaluation of Cloud-picture data has been elaborated. It must not be considered as either exhaustive or rigid. In the investigation of different problems, the Cloud-picture protocols may be analyzed in different ways. The schema offered here provides categories for the analysis and comparison of imaginative performances at different age levels and in different racial groups. The categories are sufficiently numerous and flexible to make a qualitative differentiation between the performances of the individual subjects. They are not expected, however, to diagnose the subject's "inner life." A summary of the schema and an evaluation of a single protocol are presented below.[6]

A. Content of Imagination

(1) Total number of items reported.

(2) Number of total views. (A total view is an interpretation which embraces the whole picture from any one of the four positions in which the picture may be held.) Number of items reported which are meaningful elaborations of such total views.

[6] The text of this list and the following protocol I owe again to Miss MacDonald.

(3) The number of instances in which the description of the items is elaborated in detail.

(4) The number of items which may be regarded as indices of information acquired in the child's reading or school work.

(5) The number of usual items and the number of unusual items.[7]

(6) The number of repetitive responses and the number of diverse responses. (An item may be regarded as repetitive when it is reported on two or more occasions and is not differentiated in each instance by the addition of specific details.)

Evaluations under categories 2-6 may be expressed as percentages of the total number of items reported in the protocol.

(7) Definition of the predominant directions of the content: fairy tales, geography, fire, atrocities, animals, erotic symbols, etc. Examination of the data for any general stereotype or any specific repetitive response that occurs in all of the pictures.

Note: A preliminary survey of our results indicates that the total number of items found in the Cloud pictures is less at the thirteen-year level than at the nine-year level. Although the individual differences within each age-group are greater than the mean difference between the groups, some indication is given that the onset of pubescence may be accompanied by a partial inhibition of spontaneous fantasy-expression.

B. Mode of Interpretation

Determination of the form in which the imaginative interpretation characteristically expresses itself. The categories elaborated by Struve may serve to orient this evaluation:

(1) Enumerative mode;
(2) Explanatory mode;
(3) Intuitive mode;
(4) Inventive-unregulated mode;
(5) Inventive regulated mode.

Note: These modes are empirically derived and descriptive in function. In dealing with age-levels and sociological groups other than those treated by Struve, elisions and additions may be necessary. In our experiments it was found that the enumerative and explanatory modes were often interwoven in the same interpretations. Children who enumerated disconnected items frequently added explanatory concepts and bits of academic knowledge. Similarly the intuitive and the inventive-unregulated modes were often intermingled. Children who developed vivid and plastic imagery were often carried away by uncontrolled and irrational inventions. Further investigations will determine whether combinations of these forms of interpretation are sufficiently general to necessitate a redefinition of Struve's modes.

[7] This cannot be accurately determined until more data have been collected. What is needed is a frequency table listing the frequency of occurrence of various imaginative interpretations among a fair sample of subjects.

C. The Drawing-Task

(1) Reality-boundedness: the degree of correspondence between the outline drawn by the subject and the "real object" which it represents.

(2) Form-boundedness: the degree of correspondence between the outlines drawn by the subject and the forms presented in the picture.[8]

(3) Memory: the facility with which the child is able to rediscover the items reported in the initial period of the experiment (notation of the number of times it was necessary to prompt him).

(4) Perseveration *vs.* Flexibility: the number of times, if any, that the child gave two or more imaginative interpretations for a single region of the Cloud picture; the child's ability to shift images for the same stimulus as opposed to the perseveration of the initial image.

Note: In order that marked drawing inability may not introduce an artifact into the evaluation of reality- and form-boundedness, the subject's drawing performances in the school classroom should be reviewed by the experimenter.

The category of reality-boundedness has been added to that of form-boundedness to indicate that the child may be bound in the graphic expression of his fantasy, not only by the forms present in the stimulus, but also by the stringency of his conception of the object to be drawn. Our results show that both categories bear a relation to suggestibility. Highly critical children who are pedantically aware both of the demands of the object and the demands of the stimulus-form, are generally found to react negatively in the test of suggestibility. Those who are bound in one direction only preserve a margin of flexibility within which suggestions may be received positively.

D. Suggestions

(1) Positive *vs.* Negative Suggestibility: the number of suggestions which were accepted, as a fraction of the total number of suggestions given; the average time for the acceptance or rejection of the suggestions.

(2) Sources of suggestibility.

(a) Real belief in the existence of items suggested. (Here the suggestion presents a purely intellectual task, similar to the solving of a puzzle.)

(b) Ambition or rivalry. (The experimenter's statement that other children have found the suggested item may arouse competitive impulses.)

(c) Compliance to the experimenter.

Note: An evaluation of the sources of suggestibility will represent an insightful inference based upon whatever clues the experimenter is able to gather. It is recommended that the experimenter interview personally the subject and his teachers and that observations concerning the subject's emotional reactions to the task and to the experimenter be included in the protocols. The reliability of all of the evaluations may be estimated from their consistency or lack of consistency throughout the three protocols.

[8] The category of "form-boundedness" was invented by Jacobsen, *op. cit.*

Analysis and Evaluation of a Protocol

Subject: C.C.S. Negro; male; age, ten years; I.Q. 126. Cloud Picture B.

A. Content Analysis

Total Views	Details
(1) Position: vertical 1. Title: People in a cave.	People in a cave. People outside. Boulder overhanging cave. Boulder falling. (Walls)[9]
(2) Position: horizontal 1. Title: March during a war.	Soldiers. Two commanders. Overhanging trees. Big trunks of trees. Enemy fighting. Charge. (summary statement)
(3) Position: horizontal 2. Title: Romans after Indians.	Indian shot. Roman soldier who killed him. Roman captive. Indians running away Trunk of tree. Pillar supporting roof of cave. Dying Indians.
(4) Position: vertical 2. Title: Storm in Tropics.	Dense trees. Dark clouds. Rocks. Water dashing over them.

Summary of Content:

(1) Twenty-four items were reported.
(2) Four total views were reported, one for each of the four positions in which the picture was held.
(3) Fourteen items were given descriptive details.
(4) Seven items referred to academic information. (References to "Indians," "charge," "Romans," "Tropics.")
(5) Two items were "repetitive": "trunk of trees," "cave." Each of these items occurred twice in the report. Other items which recurred were differentiated on each occasion by specifically different details.
(6) The individual items of the content were highly divergent. There appeared, however, a tendency to emphasize "military" and "nature" scenes in the total views. Both types of scenes involved the presence or threat of catastrophe in a greater or lesser degree.

[9] Parentheses indicate details added after initial report period.

H1

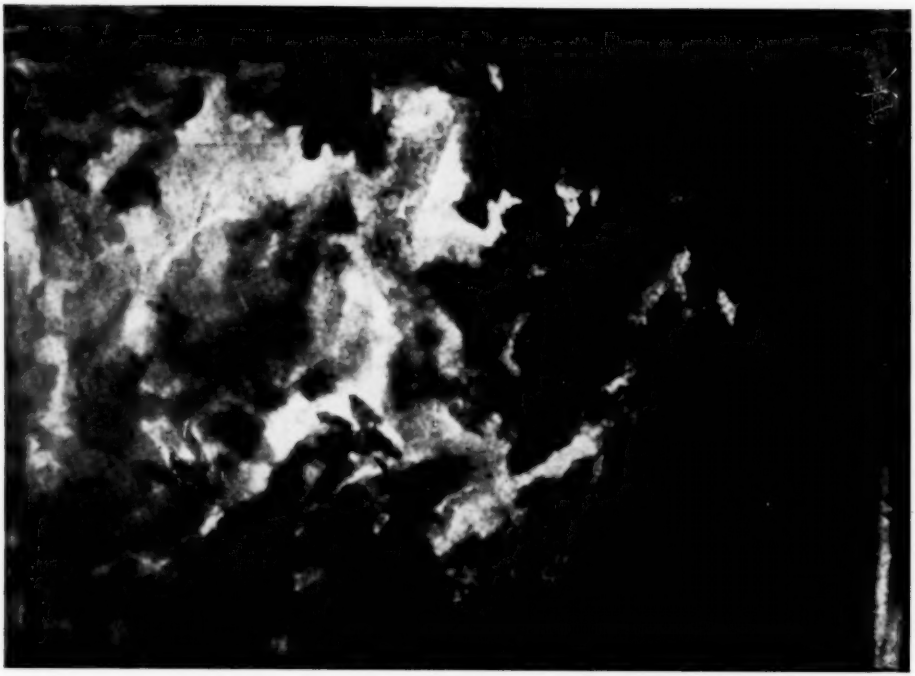
V1

V2

H2

FIGURE I. CLOUD PICTURE A

H1



V1 V2

H2

FIGURE II. CLOUD PICTURE B

H1

V1 V2

H2

FIGURE III. CLOUD PICTURE C

B. Mode of Interpretation

The subject's interpretations seem to belong to the *inventive-regulated type*. The fantasies emerged as total scenes, the partial aspects of which were logically relevant to the whole. The characteristic procedure was to note a few items, to entitle the scene and then to elaborate the details. The imaginative process was definitely controlled by a critical demand that all of the items within any one scene should be congruent. The same absence of spontaneity, the same critical planning, appeared in the subject's reaction to the drawing-task. He invariably looked at the picture for several minutes and planned the position of each item before he picked up his pencil.

C. The Drawings

A tracing of the subject's outlines for the "Romans after Indians"-scene is presented in Fig. IV. The original was done with light, precise and frequently interrupted strokes. All of the drawings were executed with painstaking care and perfect assurance.

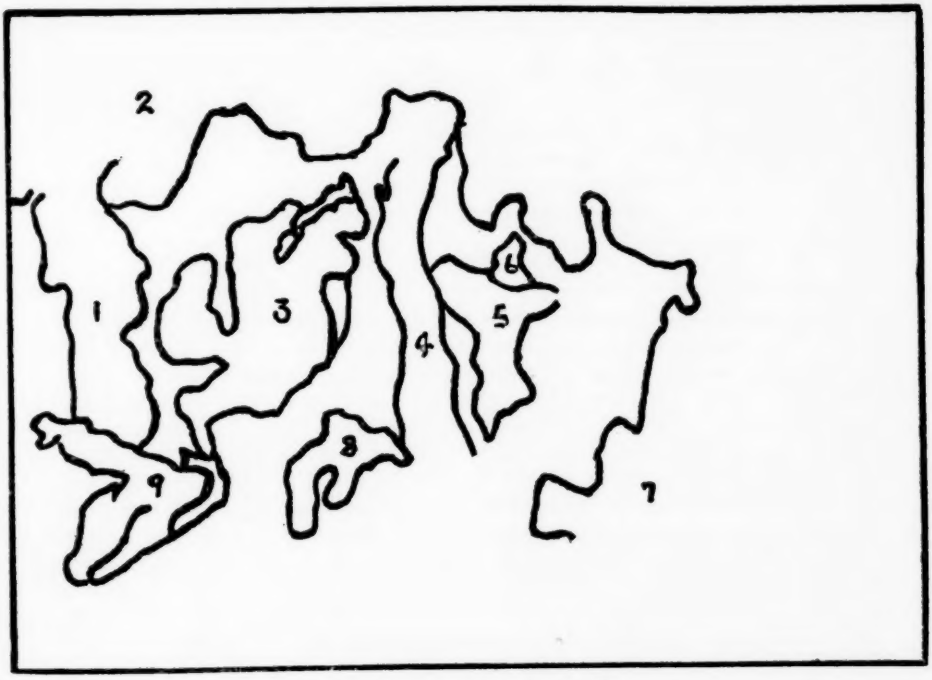


FIG. IV. CLOUD PICTURE B
Title given by subject: Romans after Indians

1. Pillar supporting the roof.
2. Roof.
3. Running Indian.
4. Trunk of a tree.
5. Roman soldier.
6. Dying Indian.
7. Wall.
8. Captive.
9. Indian dying.

(1) The reality-boundedness of the drawings is slight. (On an arbitrary scale from 1 to 5, a score of 2 may be assigned.) The outlined figures do not reproduce the objects to which they were referred with any exactness. (The human figures lack facial features, limbs, etc.) Moreover, none of the forms are rigidly and completely defined. The subject was content merely to suggest the "real objects" in his drawings.

(2) The form-boundedness of the drawings is also low (score 2-3 on an arbitrary 1-5 scale). The more salient regions of the picture were not traced precisely, and lines were superimposed upon homogeneous regions. The drawings represent the achievement of a compromise between the claims of objective reality and the claims of the stimulus-configurations.

(3) The subject's retention of items reported in the initial period of the experiment was perfect. He traced the items without any omissions and without any change in the localization initially assigned to them.

(4) The interpretations perseverated. In no instance was any region of the picture reinterpreted or given an alternative interpretation.

D. Suggestions

(1) The subject reacted positively to all three suggestions.

a) Dog (Designated in seated position. Successively completed: head, back, paws, collar on neck.) 1 min. 25 secs.

b) Bat (Delay. Question of the subject: flying or standing still? Answer: wings spread out.) 2 min. 45 secs.

c) Church with steeple. (Steeple was not completed in the drawing.) 2 min.

Average time for acceptance of suggestion: 2 min. 3 secs.

(2) The subject's suggestibility seemed to derive not only from social deference to the experimenter, but also from a strong personal ambition to excel. He accepted the experimenter's statement that other children had found a bat, and he attempted to find exactly the same bat which the others had found. His drawings of the suggested objects showed a greater approximation to reality than did his other drawings. Both the inception and expression of suggestibility seemed to be influenced by *socially competitive impulses.*

Two other protocols obtained from the same subject yield a similar picture of his imaginative processes. They, also, reveal a diverse content in which total scenes dealing with serious, nonplayful events predominate; interpretations which are verbalized in a logically controlled and systematic manner; graphic expressions of fantasy which achieve an intelligent compromise between reality- and stimulus-demands; high suggestibility following from a concentrated effort to excel.

Applications

The utilities and limitations of the new method will be determined largely by future investigations. The available data indicate, however, that the method may be used to clarify the processes and content of imagination in general, their relation to other mental

functions and their symptomatic significance for psychodiagnosis. Only intimations concerning these applications of the technique can be given in this preliminary paper.

The practical appropriateness of the method in a study of imaginative processes in general is evidenced by a comparison of Cloud-picture material with data secured through other techniques.

In our experiments, both the Rorschach Test and the Cloud pictures were employed. It was found that the Cloud pictures permitted the subjects greater freedom both in the production of the content and in the selection of the form of their interpretations. The Rorschach protocols showed a tendency toward stereotyped contents and a preponderance of wholistic interpretations. The univocal definition of the figures, achieved by saliency of the colors and symmetry of the forms, caused them to be interpreted as wholes by subjects who interpreted the Cloud pictures in a partial or enumerative fashion. The theoretical advantages contingent upon the greater scope for spontaneity afforded by the Cloud pictures have been mentioned earlier. Another advantage became evident. The three Cloud pictures were presented at intervals of one week, so that the subject was able to approach the task with a fresh impulse on each occasion. The Rorschach Test requires that the whole series of ten pictures be presented at one sitting. Although in disobeyance of this rule we divided the series into two groups of five pictures each, we often observed distinct fatigue and a tendency for the perseveration of responses.

Struve has used the Cloud test in conjunction with a linguistic test of imagination, i.e., the modified Masselon Test. This test offers three stimulus-words, and the subject's task is to combine the words meaningfully and to invent a little story about them. The great majority of subjects manifested analogous forms of behavior in the picture and word tests. The results suggest that the mode of interpretation characteristically elicited by the Cloud test is general to other forms of imaginative activity. The occasional inconsistencies which were found between performances in the two tests were especially instructive for psychodiagnosis in betraying the presence or absence of a specific limited talent.

The Cloud-picture method has also been applied to the problem of the relation between imagination and other mental functions. Comparing the performances in the Cloud test with those in tests of intelligence, no clear congruence has been found. For example, the correlation between number of items discovered in the Cloud-

picture and intelligence-test rating[10] (38 subjects) was found to be +.005. It may be supposed that children possessing a sober and acute type of intelligence, together with a predominantly critical attitude, are prevented from imagining any content that is not clearly justified by the objective conditions. Such children were frequently found to depreciate the whole task as childish. On the other hand, a rich and almost inexhaustible fantasy may be developed by children who do not occupy a high place in the intelligence rating scale. But we must suppose that the type of imaginative activity designated as the "inventive-regulated" type *needs* intelligent thinking for achieving its constructions. If, in such cases, the intelligence test yields low indices, a certain doubt must be cast upon the exhaustive value of the intelligence test score. The general conclusion seems to be, not that "imagination" and "intelligence" are two "faculties," i.e., separate forces, in the sense of the obsolete "faculty psychology," but rather that they are interdependent functions of the unique person and of his goal-directed activity. This interdependence may exhibit various aspects in one test situation or another. Under certain circumstances one function may even be stressed at the expense of the other. Nevertheless, there are certain performances which will lie beyond the capacities of low grades of intelligence. In sum, although a general correlation between numerical indices of the two test performances apparently does not exist, qualitative analysis and interpretation of individual cases seems to indicate the existence of interrelationship between the two functions.

The Cloud-picture method enables us also to approach the problem of suggestibility in a new way. In telling the subject that "other children saw this or that" we awake an ambitious desire to equal the others. This desire directs the further process of imagination and often the attempt is rewarded: the subject "sees" the suggested object. Whether a high degree of suggestibility as indicated by the Cloud test may be regarded as a symptom of a high suggestibility in general remains an open question. It would be worth while to investigate the possibility by applying other tests of suggestibility, e.g., *Aussage*-experiments, to the same subjects.

The problem of racial differences in the imagination of children has also been investigated by means of the Cloud-picture method.

The appropriateness of the material for testing different racial

[10] In our experiments the Henon-Nelson group-test of mental ability, Form A, was used.

groups is obvious, since the stimuli are so vague and ambiguous that no specific cultural conditioning is involved. It was one of the specific tasks of our experiment to investigate the question of racial differences between white and colored children. An attempt was made to isolate the racial factor by arranging pairs of children, one white and one colored, who were in all other respects (i.e., age, I.Q., social condition) as nearly analogous as possible. The results have not yet been analyzed in detail, but it is evident that the differences between the two groups cannot be expressed in simple quantitative figures. If there are racial differences which influence the reactions in addition to the very great influence of individual tendencies, these differences must lie predominantly within the realm of qualitative characteristics rather than in that of quantitative scores.

Finally, mention may be made of still another indication of the dependence of the Cloud-picture reactions upon the spontaneous tendencies and internal peculiarities of the subjects. The Cloud pictures are apparently far less dependent upon *development* than most other test performances. There are many instances in which individual tendencies seem to overshadow the progress that might be expected with increasing age. Older children may produce types of response which are much poorer and more primitive than those of younger children who surpass in the vividness and abundance of their imaginative processes.

The features of the Cloud-picture test which have already been mentioned suggest that the method may prove a promising instrument for psychodiagnostic purposes. Unfortunately, evidence to confirm this hope does not exist at the present time. It is one of the aims of this preliminary publication to submit the new method to the attention of psychoclinicians and psychiatrists. It is they who may be able to discover pathological deviations in the responses and to determine the diagnostic significance of the test. It must be emphasized, however, that the method can develop an eventual psychodiagnostic value only by being inserted into a comprehensive series of other techniques, and not by its isolated application.

For the purpose of investigating imaginative processes, I am convinced that the opportunity offered by the Cloud pictures for a study of spontaneous tendencies fully compensates for the apparent disadvantages caused by the so-called "lack of scientific exactitude." Exactitude is not, after all, a rigid concept limited

to mathematical calculus. It depends, rather, upon the careful manipulation of that method which is most appropriate to the problem to be solved. If the problem to be investigated concerns predominantly qualitative features of behavior, then the truly exact method must pay justice to these features rather than limit itself to the production of inessential scores. In this way I hope that the Cloud-picture method will develop, sooner or later, its own "exactitude."

CHARACTER TRAITS DERIVED FROM BIOGRAPHIES

FRANZISKA BAUMGARTEN

University of Bern

Since the work of G. Heymans, biographies have frequently been used for the purpose of psychological research work. Biographies may serve as a basis either for tracing the life history of a person or for finding out any special peculiarities of the mental structure.

The question may be raised as to how far biographies are to be relied upon as a source for scientific research work. For instance, it struck me that the biographies of some prominent personalities—well known to me—do not stress those characteristic traits which I myself was able to perceive. My doubts are intensified when I myself try to make use of biographies in a characterological problem on which I am at present working. I can easily conceive that the same person shows quite different characteristics according to the descriptions given by different people. Here is an especially striking example: Luise Kautsky describes the well-known socialist, Rosa Luxemburg, as being endowed with extremely positive qualities, such as generosity, charity, compassion, kindness, readiness to make sacrifices, idealism, and so on. On the other hand, biographical notices in the conservative press, characterize her as a monster, a "hyena," and a murderess.

However, apart from such extreme cases, even authors who are psychologically interested and emphatic in their impartiality, differ in their designations of characteristics of the same person. As an example, we may quote some of last year's publications, that is to say, the biographies of Mary Stuart. These were written, one following shortly after the other, by Stephan Zweig and Marjorie Bowen. The characteristics mentioned in both books were taken by us and compared. In doing so, we only took into consideration the individual character qualities, but we refrained from doing so where

Bowen—as she often did—failed to express these clearly, using instead lengthy and roundabout phrases, as, for instance, "unworthy treatment of the second husband" or "kind treatment of the man in order to deliver him to his enemies." Although a number of sentences indicate qualities such as self-interest, deceitfulness, faithlessness, etc., we omitted to explain them in detail. Interpretations of this nature might be put down to a matter of individual opinion. Accordingly, we only took out those characteristics which stood out clearly in both books.

The list of qualities—which we have not been able to give in extenso here—shows that Zweig attributes 92 and Bowen 119 to Mary Stuart. Only in 22 cases are the same qualities mentioned and even these are often expressed synonymously. To give examples: constancy-firmness, deceitful-mendacious. This means that only one-fifth of the qualities attributed by Bowen to the unhappy Queen, are also to be found in Zweig's book. On the other hand, for the completeness of the portrait, both authors have made mention of numerous qualities of the greatest importance. Zweig alone speaks about firmness, constancy, reserve, and helpfulness. Sometimes the two authors even mention opposite characteristics. Zweig writes of truthfulness, sincerity; but Bowen affirms mendacity. Zweig denotes timidity; Bowen, fearlessness. We are not here concerned with opposite qualities as reported by the same author, such as cautiousness-incautiousness, since this deals with another problem, which is that of the opposite qualities in the structure of personality.

The point with which we do concern ourselves is the distribution of the characteristics according to their kind, as we find them set out by the two authors in describing one and the same person, Mary Stuart.

In order to serve our purpose, we have ranked the characteristics in the following classes:

(a) positive, such as honesty, resoluteness;
(b) negative, such as deceitfulness, stubbornness;
(c) neutral, which may be positive as well as negative, according to the circumstances in which they are used. It may, for instance, depend upon the aim of the ambition as to whether ambitiousness is a positive or a negative quality.

By means of this classification, we arrive at the following proportional percentages:

	According to Zweig	*According to Bowen*
	Per Cent	Per Cent
Positive qualities	54.23	26.86
Negative qualities	35.86	67.22
Neutral qualities	9.78	5.88

The proportional relation is:

According to Zweig: 3 : 2 : 0.5.
According to Bowen: 1 : 2.5 : 0.2.

This shows how the positive qualities prevail in Zweig's biography, and in Bowen's case the negative qualities. Thus the Queen's "psychic profile" is represented quite differently by the two authors.

As a result of these psychological tracings based on comparison, we find two fundamentally different conceptions of the same person. The outcome of psychological research work would, consequently, depend entirely on which of the two biographies is chosen as a basis. This involves the possibility of arriving at two absolutely different results. The question now arises, with which all those who make use of biographies in their research work are concerned, as to which of these two is the correct one. Is Zweig's the more reliable or Bowen's? Where lies the choice? This would probably depend on mere subjective reasons, or again on the fact that the research worker is familiar with only one of the two books.

What do we learn from this difficulty, which is bound to occur with biographies of nearly every prominent personality?

(1) That "free characteristics" given in biographies of a literary kind are always subjective;
(2) that research workers who base their work on biographical statements must, therefore, be extremely careful in the application thereof;
(3) that such an individual conception of a person is obviously favored by the fact that the science of characterology has not so far laid down any schedule of characteristics which may be used to describe a person without any ambiguity.

Of course, a biographer is entirely at liberty to work in his own way and not tied down to any fixed plan of characterization. This license not only makes it possible for authors to describe the same person from their own point of view, but also to stress quite different qualities in painting the picture. It is, however, vitally essential that psychologists should come to an agreement as to which qualities are useful in describing a personality. At the same time, such a schedule could serve as a guide for biographers in describing a character.

TRENDS IN THE STUDY OF PERSONALITY

(As Revealed in the 1937 Meeting of the American Psychological Association)

DONALD W. MACKINNON

Bryn Mawr College

It is customary to look upon the program of the annual meeting of the American Psychological Association as a report upon the current state of psychology in the United States and to seek in an analysis of the relative number of papers in the various fields some indication of the shifting trends of psychological interest. An analysis of the papers read at the 1937 meeting of the association held in Minneapolis early in September reveals a continued widespread interest in the psychology of personality. More significantly it reveals the development of certain new trends in this field.[1]

One of the most important of these trends appears to be an increasing emphasis upon the experimental investigation of problems basic to personality. The meeting for the first time of a section on "Conflict and Frustration" gave opportunity for presentation of a number of experimental studies on the dynamics of behavior. Emphasis has been placed recently upon the role of frustration in the development and structurization of the personality and upon the hypothesis that personality traits depend upon the type of restructurization of the field which follows the blocking of needs.[2] Accordingly, these experimental investigations of the effects of conflict and frustration are particularly significant for the psychology of personality.

R. Barker, T. Dembo, and K. Lewin reported observations of

[1] The present report stresses what the writer considers to be the significant novel developments in the study of personality and represents an individual evaluation. Contributions here neglected may well in the long run be judged the more important. They are omitted from discussion not because they are considered unimportant but merely because they do not fit into any of the trends with which the present paper is concerned. An abstract of all papers presented at the meeting can be found in the *Psychol. Bull.*, 1937.

[2] J. F. Brown, *Psychology and the Social Order* (New York, 1936), p. 240.

the behavior of thirty children, two to five years of age, in two situations which, physically identical, were psychologically different, the one having the meaning of a satisfying play situation, the other the meaning of a frustration situation. The development of a scale of age norms for constructiveness of play made possible a quantitative expression of the degree of regression under conditions of frustration—a mean regression in the constructiveness of play equivalent to twenty-two months at the 4.5 year level and to five months at the 2.5 year level being reported. Among other observations one of particular interest was that while in general, the amount of regression was positively related to the degree of frustration, at times a weaker frustration led to a slight increase of the constructiveness of play. This suggests the existence of a critical point on the frustration continuum below which frustration leads to a psychobiologically adaptive restructurization of the field and above which frustration leads to excessive distortion of the field and a less adequate response on the part of the individual. Such a critical point on the frustration continuum has been conceptualized as "frustration tolerance."[3]

An investigation pointing to some of the factors determining frustration tolerance was reported by G. M. Haslerud. Chimpanzees were trained to pull an endless belt on which, after 48 cm. had been pulled, there appeared from behind a cloth screen a food carriage containing an orange. The effects of three kinds of frustration upon this behavior were then observed. The frustrations "(a) After standard effort, empty food container appears, (b) After standard effort, wooden screen lowered to interrupt pulling, (c) Unlimited pull with no reward" were not equally potent. The interrupted pull constituted for most animals the strongest frustration, the appearance of the empty food container the weakest frustration. With unlimited pull with no reward the animals' effort and time far exceeded the standard. Particularly interesting were the observations of qualitative differences in the type of reaction to the various degrees of frustration and the increase of frustration tolerance with age.

Other investigations of conflict and frustration revealed another significant trend in the study of personality, namely, an effort to co-ordinate systematic positions within psychology. Two papers

[3] S. Rosenzweig, Frustration as a co-ordinating concept in experimental psychology, published in the *Proceedings of the N. R. C. Conference on Experimental Neuroses*, April, 1937.

reported investigations of the types of conflict described by Lewin, the results of which could be interpreted either in terms of a topological and vector psychology or in stimulus-response terms and the goal-gradient hypothesis of Hull, the latter being the preferred explanation of the investigators. C. I. Hovland, investigating differences in resolution of approach-approach and avoidance-avoidance conflicts (types 1 and 3 for Lewin), trained subjects in one experiment to draw a line from a point equidistant from two lights to the one which was illuminated and found that subsequent simultaneous flashing of both lights produced little conflict, most subjects drawing a line to one or the other of the lights. In a second experiment subjects were trained to draw a line away from the illuminated light. Here simultaneous flashing of both lights produced conflict and partial compromise movements. R. R. Sears studied Lewin's second type of conflict, i.e., between approach and avoidance responses, by training subjects to draw a line "*toward* the flash of a green light and *away* from that of an adjacent red light" and found that subsequent simultaneous flashing of both lights produced conflict with frequent blocking of the manual reaction, the resolution of the conflict occurring usually on a verbal level. Since in this type of conflict situation Lewin deduces a tendency to "go out of the field" it was unfortunate that Sears's experiment did not permit the subject to go out of the field in the manual response. These two papers indicate a most important trend in the study of personality—a sympathetic rapprochement between two different conceptual formulations—for only on the basis of such co-operative research can the relative merits of opposed theories be decided.

A panel discussion of "The Experimental Approach to Psychoanalytic Concepts" in which a psychoanalyst met with psychologists to discuss the possibilities of co-operative research between experimental psychology and clinical psychoanalysis demonstrated even more strikingly the efforts being made for a more fruitful co-ordination of different approaches to the psychology of personality. In retrospect it would appear that the purposes of the discussion would have been better served by a round table than by a panel, yet in spite of its shortcomings the panel discussion did succeed in appraising, in part, the experimental and analytical approaches to problems of personality and in formulating some modifications of both approaches which should lead to the development of techniques more adequate than any previously used for the experimental investigation of psychoanalytic concepts.

Of the many signs of the fruitful cross-fertilization of fields none, however, was so striking as the fusion of animal psychology and psychoanalysis. Neal Miller reported experiments with rats which had been suggested by Freudian generalizations, and presented a film entitled "Reaction Formation in Rats: An Experimental Analog for a Freudian Phenomenon," showing the arousal of an approach tendency, the completion of which had previously led to shock, being inhibited by the expression of a clear avoidance response. O. H. Mowrer also reported reaction formation in rats, although his paper was primarily concerned with the demonstration of "regression" in rats. Although experiments such as these are still somewhat distant from the deeper conflicts presumably playing a role in human psychopathology, they have real importance in offering an opportunity for an experimental investigation of those psychological processes and mechanisms which have been described by the psychoanalysts and which are so tremendously important for the psychology of personality.

While the psychology of personality has advanced along the new fronts just described there has been no slackening of activity upon its older battlefields. The specificity-generality controversy goes on. Specificists now have V. E. Herrick's report that the success and failure attitudes developed in his monkeys "were specific and exerted little, if any, influence over objects and situations not connected with their origin," and A. T. Jersild's report of "inconsistencies between a child's defense of a general attitude and his specific prejudices; occurrence of 'mature' and 'juvenile' forms of reasoning in the same child, depending, apparently on his information and experience in connection with specific topics." Generalists interested in the consistency of personality have W. C. Olson's report of a marked consistency in the social adjustment of children over a period of six years, while those interested in problems of style and idiom have evidence for a style of expression extending beyond the individual to the nation in C. L. Stone's report that when musically unselected students were asked to identify the composer, composition, and nationality of each composition, their percentage of correct identifications for each nationality was considerably better than chance when the names of composer and composition were unknown.

The study of attitudes continues to be an important concern of the psychologist interested in personality. S. W. Cook and E. Raskin reported a further validation of Stagner's "Attitude Toward

Fascism" scale. Using an "autobiographical schedule" with groups of professed radicals and non-radicals and examining the data for significant differences between the groups, M. H. Krout reported a list of early experiences and personality traits characteristic of radicals. C. R. Pace reported an attempt to measure liberal attitudes by asking S's to check in which of five specific ways they would behave in thirty current "social-political-economic situations." Correlating scores on this attitude test with scores on an information test he found positive correlations between liberal attitude and information. Trenchant criticisms of the usual type of attitude scale were made by R. H. Ojemann and J. L. Graham, both of whom showed that the vagueness of the concepts in most scales result in an inaccurate determination of the testee's attitude. Ojemann suggested a more precise definition of the "key-concept" and less simple statements to decrease the opportunity for verbalizing and showed that a scale so constructed yielded data that correlated reasonably high with behavior under known conditions. Graham proposed a method of breaking up attitudes and reported data indicating that "systematic analysis of the component parts of an attitude yields more exact measurement than a single score or rating."

Personality inventories and adjustment questionnaires received their usual share of attention. When one is faced with the responsibility of practical action in counselling and guidance, inventories and questionnaires yield little of value as compared with the interview according to Kimball Young. On the other hand, E. A. Gaw reported the usefulness of the Ohio Check List in consultation with students. The analysis of a single test of "self-reliance" into factors tentatively named "independence of decision in personal matters," "social initiative," resourcefulness in the use of spare time," and "personal responsibility" was reported by L. H. Stott. Of the analyses of adjustment questionnaires, the most searching was reported by Emma Layman who analyzed 782 personality test items from 16 questionnaires. Reducing the number of items finally to 67 and submitting them to factor analysis she found at least 12 factors involved in the final reduced test. Thus, factor analysis continues to be an important tool for the psychologist interested in personality, and papers dealing with some of its more technical problems were read by R. C. Tyron and Paul Horst.

The problems of personality development have long been investigated. That an extensive integration of techniques in this field is being made was revealed in H. E. Jones's report of the study of

adolescence being made at the Institute of Child Welfare. His theoretical discussion of methodological problems was given concrete meaning in the presentation by W. J. Cameron of an illustrative case study of growth in adolescence. This paper was unique in being, so far as the writer knows, the only presentation in the meetings of a picture of a single total personality. In contrast to this single case study was G. K. Bennett's report of the interests and activities of three thousand adolescents.

If the papers presented at the meeting of the American Psychological Association are representative of the interests finding expression in research in any year, the most important current trends in the field of personality appear to be an increasing emphasis upon the experimental investigation of basic problems and attempts at co-ordination both within psychology and between psychology and other sciences in the study of personality. But while this work in new areas of the field is what is most striking today, it neither obscures nor lessens the continuing interest and important research in those areas which have been longer worked.

BOOK REVIEWS

THE EVOLUTION OF MODERN PSYCHOLOGY. By Richard Müller-Freienfels. Translated by W. Beran Wolfe, M.D. New Haven: Yale University Press, 1935. xvi, 513 pp. $5.00.

Richard Müller-Freienfels has long been known to students of German philosophy as a stimulating writer, but in spite of his productivity there has been no work of his accessible to English readers until the Yale University Press commissioned the late Beran Wolfe to translate the work under review. The translation was made from the typescript, which would indicate that the work was first published in English. *The Evolution of Modern Psychology* may be regarded as a history, and is more in the historical style than the various histories of psychology which we have to date.

It would be futile in the compass of a brief review to present the substance of work so comprehensive in scope. The chapter headings would not give a fair notion of what the volume covers.

Starting out with a historical introduction in which the author exhibits his breadth of learning, he proceeds to show how psychology became conscious-minded. The various early issues of sensationalism, the elements, the problem of localization, the question of imagery, the continual conflicts between the analytic and the synthetic schools are treated with great clearness and aplomb. The author's æsthetic background and learnings have assisted him considerably in enlivening topics which have elsewhere found dull exponents.

Is is evident that Müller-Freienfels is at home in most of the fields he discusses, and his own position seems to be more discernable when he speaks of other people's views than when he presents it directly. Although he does not mention the kinship, it would, I think, be appropriate to class him with Mary Calkins under the rubric of self-psychologists. He is continually stressing the totalitarian point of view in psychology which may be symbolic to him; we should prefer to see that word restricted to the sphere of current political situations.

The chapters on the unconscious and on collective psychology are probably the best in the book, while the appendix on parapsychology is an open-minded survey of all the phenomena usually labeled Psychical Research—open-minded yet conservatively sensible and balanced.

The weakest section in the book is that on characterology. From Paulhan, who wrote his book nearly fifty years ago, he jumps right into the Germany of Klages and leaves off with his work published in 1927. The survey, even of Klages, is skimpy, and no attempt is made to draw the line between characterology and psychodiagnostics. The problem of typology appears in jets almost everywhere, whereas it should have received a chapter for itself.

On the whole, the author's perspective is unexceptionable. He has

an eye for relationships, but is inclined to do a bit of meandering, sometimes introducing a psychologist in various connections, to such an extent that he may appear to partake of opposing systems. James, who, he claims, has influenced him greatly, is thus found among the "kings" and the "prophets." Müller-Freienfels has done more justice to American psychologists than have most of his compatriots, but there is still some territory which appears to have been untouched by him. Behaviorism has received twice as lengthy an exposition at the hands of Müller-Freienfels as Russian psycho-reflexology. It is manifest that he wished to be especially fair to Watson, perhaps because it is here that the author swerves from his judicial temper and reveals his anti-mechanistic attitude. He writes: "I have given a thorough behavioristic exposition of the ideological trend of behaviorism, for the most part in Watson's own words. And I have done this without irony, something which the behaviorist incidentally, has scarcely any right to understand, according to his own principles; for in irony, speaking and thinking are not identical but are antagonistic. At any rate, it is obvious that behaviorism, in spite of all its eloquent scientific pretensions, ultimately resolves itself into a matter of prophecy. Of all the paradises promised us by clerics and philosophers, the utopia of behaviorism impresses me as the most boring, because it transforms the universe into a machine shop."

It is fortunate that the translator has been able to get into the spirit of the work so as to preserve the felicitous phrases and picturesque language of the author, but it is a pity that he has deemed it advisable to translate the titles of German works that had never appeared in English, thus in some cases at least, leading to confusion, for the student would need to retranslate into German when looking up a reference. The bibliography of some three hundred works, however, will be of some help toward this end. The volume has been carefully prepared and the usual excellence of production for which the Yale University Press is known has been maintained.

A. A. Roback.

The Power to Love: *A Psychic and Physiologic Study of Regeneration.* By Edwin W. Hirsch. New York: Alfred A. Knopf, 1934. xi, 363, ix pp. $4.00.

What distinguishes this book from many others on the subject of sex is that the author seems to know his subject at first hand and, furthermore, goes into a wealth of detail in his descriptions. Even the tyro can gather that the author is conversant with the facts of sex and is critical of many of the accepted beliefs, e. g., as the depressant effects of saltpeter.

It has become almost an impossibility to bring out in a book related to the love life anything which is new, so that one book seems to be just a different version of another, but Dr. Hirsch in his presentation gives a systematic account of the numerous processes, drawing upon both the literature of the past and recent psychoanalytic material. There is a sane approach to many of the problems that have beset man and woman in this sphere of life. It is a sign of good sense when he draws the line between the rat and man, and shows the futility of such experiments of Dr. Lydston's. "The

original experimenters argued that if gland transplants were highly effective in rats then human beings should likewise be benefited. Ideas such as this are highly contagious and this one took soil in the brain of Dr. G. Frank Lydston, who decided to transplant a human testis into his own scrotum."

It is in the latter part of the book that the author begins to veer from his praiseworthy sanity. It is difficult to believe that some people are such morons in matters sexual, that a slow-motion picture must be produced and each fraction of a move must be taught, thus outdoing the Kama-Sutra of the Hindus.

One should suppose that aside from removing the prejudices and sham connected with the sex act, it would be best to let each man and woman discover the subtleties of the technique for themselves. Dr. Hirsch's very specific and detailed instructions seem not only a bit droll, but in view of individual differences, may produce unsatisfactory results. It is ludricrous to think that any individual will at such times consult a book and follow out the directions in all detail.

The Power to Love is an unusually frank book, and yet it gives an impression of earnestness. The mode of handling impotents is discussed with a good deal of understanding.

A. A. Roback.

The Growing Child and His Problems. Edited by Emanuel Miller. Contributions by *Several Collaborators.* London: Kegan Paul, Trench, Trubner, 1937. 231 pp. 6/-.

The purpose of this book is to demonstrate how mild disorders of psychological origin can be satisfactorily prevented or alleviated. It is intended for the general reader and the specialist, but more especially for the former.

Such a book should be lucid and informative, and contain sufficient positive and practical suggestions. It should not leave the lay reader with a lingering anxiety that all the children with whom he may come in contact are potential "problem" children, and that almost anything he may do is liable to aggravate the problem. It should enable the parent, the teacher, the school nurse, the health visitor, and the club leader to "gain some insight into those red lights in behavior which suggest that there is somewhere a hold-up in the mental traffic" as Dr. Miller so aptly expresses it, and be prepared to refer for psychological assistance when usual methods have failed.

This is a stern criterion, but this book has succeeded in large measure in satisfying it. Miss Chester's account of the child's "play needs" contains many valuable practical suggestions not often appreciated by teachers and parents, and she lays special stress on the right adult attitude to the child's play. To the specialist her analysis of play does not go far enough—one would like a more complete statement of her reasons for the choice of play material, but this would be outside the scope of the present book.

The detailed discussion on the problems of the adolescent boy and girl should be of considerable help to anyone who comes in contact with this age-group.

Dr. Burns takes into account the multiplicity of factors likely to cause or aggravate neurosis in school

children, and his contribution to the book should be illuminating to the teacher in that he shows how her treatment of the child may be influenced by her own unconscious motives.

Dr. Paterson Brown in his chapter on "Habits" gives the most dynamic and fundamental account of the causes of difficulties in young children and of the effect of parental attitudes. His treatment of the subject is, however, perhaps too specialized for the general reader. I think it is also likely that from a cursory reading of Dr. Clifford Allen's account of Personality Deviations, ordinary parents may become too much aware of possible mistakes in the handling of the child, causes them to lose confidence and exaggerate minor difficulties into major ones.

Miss Simmins gives a very clear exposition of the value of scientific educational guidance of the school child, which should be helpful to both parents and teachers. It is perhaps unfortunate that her account seems so disconnected with the matter contributed by the other writers. To my mind the point of view of the Educational Psychologist is too often too far divorced from that of the Mental Psychologist. Sometimes a closer co-operation between these two experts in the investigation of causative factors would be beneficial.

The general plan of the book—the wide variety of problems dealt with, and the difference in each writer's approach to them—proves stimulating reading, and should result in a more understanding and tolerant attitude towards children.

AGATHA H. BOWLEY.

INTRODUCTION TO THE RORSCHACH METHOD. By S. A. Beck. Monograph No. 1 of the American Orthopsychiatric Association. 1937. 278 pp.

This volume is a manual of Rorschach Test procedure which incorporates recent advances in technique and illustrates the particular practice of the writer. It is intended to be helpful to the neophyte and at the same time to offer a normative, objectively stable standard of procedure as a starting point for the accumulation of comparable material by different workers in the field.

The bulk of the volume is given over to the recording and elaboration of fifty-nine cases in ten groups: healthy adults of superior intelligence, feeble-minded, depressed states, hypomanic conditions, schizophrenics, neurotics, conduct disorder in adults, problem children, "inadequate test records" (exemplifying apparent misdiagnosis), and "mental hygiene cases" in which the Rorschach findings indicate the feasibility of mental hygiene efforts. The remaining third of the discussion deals with experimental technique, scoring, and a summary of the psychological processes supposed to be indicated by the several "Rorschach factors." Under scoring the contributions of especial interest are the introduction of the "organization" (Z) response, with detailed directions for scoring; the elimination of the "original" (O) response; introduction of normative tables for scoring "detail" (D) and "rare detail" (Dr) responses, and a similar set of tables for "good form" (F+) norms. Under interpretation the category of "experience balance" has been substituted for *Erlebnistypus*, and a good deal less weight

has been put upon this category than under the older procedure. Finally a table for classification of content is included as a basis for taking variations of interest into account, or, as Beck expresses it, as a means for analyzing "social trend" as opposed to "psychic structure."

Among the many positive contributions of the book, however, one serious drawback must be noted. In connection with the compilation of normative tables, no serious attempt has been made to offer a clear statement of the criteria which may be used as a basis for the classification of responses which do not appear upon the tables. For example, the determination of an F response is ambiguous in that it apparently depends not only upon correspondence with stimulus configuration but upon the frequency with which the particular stimulus feature receives the given response content. Thus "interest trend" is seen to be intimately associated with the determination of a factor in "psychic structure." A similar situation obtains in regard to the scoring of movement (M) responses. The omission of theoretical discussion of interpretations is understandable (though regrettable) in consideration of the particular aims of the book, but lack of "theoretical" discussion of criteria in scoring cannot pass without comment.

S. H. MacColl.
Duke University.

HUMAN GENETICS AND ITS SOCIAL IMPORT. By S. J. Holmes. New York: McGraw-Hill Book Co., 1936. vi, 414 pp. $3.50.

Professor Holmes is particularly well qualified to write on a subject which deals with application of biological principles to Human Welfare and in this book he has given a very clear and scholarly account of the relation of modern-day genetics to social progress. The first eight chapters deal with hereditary mechanics; and while some knowledge of fundamental biological principles is required of the reader, the facts and theories of inheritance are treated in a nontechnical manner. The latter chapters deal with such subjects as heredity in disease; differential birth rates; effects of migration, inbreeding, war, population growth, etc. The final chapter contains proposed measures for Race Betterment, a subject on which Professor Holmes has written before so ably and judicially. At the end of each chapter the author has arranged a few suggested readings and a series of questions which should stimulate the serious reader and give him a clearer insight into the subjects treated. The book is well illustrated with photographs, charts, and tables. Professor Holmes believes that social problems can ony be solved by a study of biological facts and that only through education can we hope to have a better race.

F. G. Hall.
Duke University.

THE PSYCHOLOGY OF A SUPPRESSED PEOPLE. By J. C. Heinrich. London: Allen & Unwin, 1937. viii, 144 pp. 3/6 net.

The small compass of this book does not prevent it from containing a great deal of material that will be useful to two classes of people. Social psychologists will find it a little mine of useful case-material to cite. Missionaries and other persons whose work lies in India or with colored peoples, will find it an extremely valuable aid, because of the excellent common sense sharpened by observation and true psychological insight of its author.

Familiarity with much of the present output of psychological books written by clergymen has made the present reviewer approach a new one of this class with skeptical trepidation. In the present instance, however, this proved quite unjustified. The author obviously prepared himself for his work by assimilating considerable academic psychology, negro biography, etc. These he had digested thoroughly, and produces his knowledge without either parade or maladroitness, to enrich and to help him organize his very useful original memoranda of personal experiences in the Indian missionary field.

PRYNS HOPKINS.

NEW FRONTIERS OF THE MIND. By J. B. Rhine. New York: Farrar and Rhinehart, 1937. 274 pp. $2.50.

New Frontiers of the Mind presents in popular style and in narrative form an account of the experiments in extra-sensory perception (telepathy and clairvoyance) carried out by Dr. Rhine and his collaborators. Since a more complete account of the methods used and of many of the results described has been presented in *Extra-Sensory Perception* by the same author (reviewed earlier in this Journal) only the new topics discussed require mention here. One chapter is devoted to a discussion of work done in this field by investigators in other laboratories. A majority of those reported were confirmatory of the phenomenon of extra-sensory perception. The general psychological conditions favoring the appearance of the phenomena are described. Recognition of these conditions is regarded as essential in serious experimentation in this field. In the continued analysis of both favorable and unfavorable conditions lies the hope of bringing the phenomena into closer relationship with recognized psychological processes—one of Dr. Rhine's present aims. In a discussion of its physical relationships it is concluded that the "processes of extra-sensory perception bear no discoverable relationship to physics."

Perhaps the experiments, here first discussed, on the relationship of these phenomena to time (precognition, retrocognition) offer one of the greatest potential contributions of this line of work to our knowledge of the nature of human personality and its relation to physical and biological processes. Some of the difficult problems of experimentation and interpretation are briefly touched upon in the final two chapters. It is to be expected that more detailed presentations of these experiments will shortly appear in the newly founded *Journal of Parapsychology*, to which those seriously interested in this whole field may turn for reports of current research.

KARL ZENER.

Books Received

PSYCHOLOGY OF PERSONALITY. By Ross Stagner. New York: McGraw-Hill Publishing Company, 1937. xi, 453 pp. with index. $3.50.

PERSONAL MEMOIRS OF H. P. BLAVATSKY. Compiled by Mary K. Neff. New York: E. P. Dutton & Company, 1937. 316 pp. with index. $3.75.

THE REAL USE OF THE UNCONSCIOUS. By Chandler Bennett. New York: The Dial Press, Inc., 1937. vii, 368 pp. with index. $3.00.

TRY LIVING. By William Moulton Marston. New York: Thomas Y. Crowell Publishing Company, 1937. 228 pp. $1.75.

ART AND PRUDENCE: *A Study in Practical Philosophy.* By Mortimer J. Adler. New York, Toronto: Longmans, Green and Company, 1937. xiv, 676 pp. with index. $5.00.

WHAT MAN HAS MADE OF MAN: *A Study of the Consequences of Platonism and Positivism in Psychology.* By Mortimer J. Adler. With an introduction by Dr. Franz Alexander. New York: Longmans, Green and Company, 1937. xix, 244 pp. $3.50.

COUNSELING TECHNICS IN COLLEGE AND SECONDARY SCHOOL. By Ruth Strang. New York and London: Harper & Brothers, 1937. x, 153 pp. with index. $2.50.

ACHIEVING CHRISTIAN CHARACTER. By Frances Cole McLester. Nashville: Cokesbury Press, 1937. 160 pp. $ .60.

BEHAVIOR AND BACKGROUND OF STUDENTS IN COLLEGE AND SECONDARY SCHOOL. By Ruth Strang. New York and London: Harper & Brothers, 1937. xiv, 499 pp. with index. $4.00.

THE QUESTIONING MIND. By Rupert Clendon Lodge. New York: E. P. Dutton & Company, Inc., 1937. vii, 311 pp. with index. $2.75.

SURPRISE AND THE PSYCHOANALYST: *A Study of the Conjecture and Comprehension of Unconscious Processes.* By Dr. Theodor Reik. New York: E. P. Dutton & Company, 1937. vii, 287 pp. with index. $3.50.

PERSONALITY DEVELOPMENT IN CHILDREN: *A Multiple Approach to a Complex Problem.* By Ernest J. Chave. Chicago: The University of Chicago Press, 1937. xiii, 348 pp. with index. $2.50.

PERSONALITY: *A Psychological Interpretation.* By Gordon W. Allport. New York: Henry Holt & Co., 1937. x, 566 pp. with index. $3.50.